



# HHS Public Access

Author manuscript

*Ann Intern Med.* Author manuscript; available in PMC 2017 February 16.

Published in final edited form as:

*Ann Intern Med.* 2016 February 16; 164(4): 268–278. doi:10.7326/M15-1789.

## Supplemental Screening for Breast Cancer in Women with Dense Breasts: A Systematic Review for the U.S. Preventive Services Task Force

**Joy Melnikow, MD, MPH,**

Department of Family and Community Medicine, Center for Healthcare Policy and Research, University of California Davis, 4860 Y Street, Suite 2320, Sacramento, CA 95817, 916-734-3160

**Joshua J. Fenton, MD, MPH,**

Department of Family and Community Medicine, Center for Healthcare Policy and Research, University of California Davis

**Evelyn P. Whitlock, MD, MPH,**

Kaiser Permanente Research Affiliates, NW

**Diana L. Miglioretti, PhD,**

Department of Public Health Sciences, Center for Healthcare Policy and Research, University of California Davis and, Group Health Research Institute, Group Health Cooperative, Seattle, WA

**Meghan S. Weyrich, MPH,**

Center for Healthcare Policy and Research, University of California Davis

**Jamie H. Thompson, MPH, and**

Kaiser Permanente Research Affiliates, NW

**Kunal Shah**

---

**Current Author Addresses:** EP Whitlock, JH Thompson: Kaiser Permanente Center for Health Research, 3800 North Interstate Avenue, Portland, OR 97227

J Melnikow, JJ Fenton, MS Weyrich: Center for Healthcare Policy and Research, University of California Davis, 2103 Stockton Blvd, Sacramento, CA 95817

DL Miglioretti, Department of Public Health Sciences, UC Davis School of Medicine One Shields Ave., Med Sci 1C, Room 145, Davis, CA 95616

K Shah, Columbia University, 6380 Lerner Hall, 2920 Broadway New York, NY 10027

**Note:** This review was conducted by the Kaiser Permanente Research Affiliates Evidence-based Practice Center with the UC Davis Center for Healthcare Policy and Research under contract to AHRQ. AHRQ staff provided oversight for the project and assisted in the external review of the companion draft evidence synthesis. The analytic framework, review questions, and methods for locating and qualifying evidence were posted on the USPSTF website for public comment prior to commencing the review; final versions reflect public input. The authors of this report are responsible for its content, including any clinical treatment recommendations. No statement in this article should be construed as an official position of AHRQ or the U.S. Department of Health and Human Services.

**Potential Conflicts of Interest:** Disclosures can be viewed at XXX

**Request for Single Reprints:** Reprints are available from the AHRQ Web site ([www.ahrq.gov](http://www.ahrq.gov))

**Author Contributions:**

Conception and design: Melnikow, Fenton, Whitlock, Miglioretti

Analysis and interpretation of data: Drafting of the article: Melnikow, Fenton, Whitlock, Weyrich, Thompson, Shah

Critical revision of the article for important intellectual content: Melnikow, Fenton, Whitlock, Miglioretti, Weyrich, Thompson, Shah

Final approval of the article: Melnikow, Fenton, Whitlock, Miglioretti, Weyrich, Thompson, Shah

Statistical expertise: Miglioretti

Obtaining of funding: Whitlock

Administrative, technical, or logistic support: Weyrich, Thompson, Miglioretti

Collection and assembly of data: Melnikow, Fenton, Whitlock, Miglioretti, Weyrich, Thompson, Shah

Columbia University

Joy Melnikow: jamelnikow@ucdavis.edu

## Abstract

**Background**—Screening mammography has lower sensitivity and specificity in women with dense breasts, who experience higher breast cancer risk.

**Purpose**—Systematic review of: reproducibility of BI-RADS density categorization; test performance and clinical outcomes of supplemental screening with breast ultrasound, magnetic resonance imaging (MRI), and digital breast tomosynthesis (DBT) in women with dense breasts and negative mammography.

**Data Sources**—MEDLINE, PubMed, Embase, and Cochrane January 2000–July 2015.

**Study Selection**—Studies reporting BI-RADS density reproducibility or supplemental screening results for women with dense breasts.

**Data Extraction**—Quality assessment and abstraction of twenty-four studies from seven countries; six were good quality.

**Data Synthesis**—Three good-quality studies reported reproducibility of BI-RADS density; 13–19% of women were re-categorized between “dense” and “non-dense” at subsequent screening. Two good-quality studies reported ultrasound sensitivity for women with negative mammography ranging from 80–83%; specificity 86–94%; and positive predictive value (PPV) 3–8%. MRI sensitivity ranged from 75–100%, specificity 78–94%, and PPV 3–33% (3 studies). Ultrasound additional cancer detection rates were 4.4 per 1,000 exams (89–93% invasive); recall rates were 14%. MRI detected 3.5–28.6 additional cancers per 1,000 exams (34–86% invasive); recall rates were 12–24%. DBT cancer detection rates increased by 1.4–2.5 per 1000 exams compared to mammography alone (3 studies). Recall rates ranged from 7–11%, compared to 7–17% with mammography alone. No studies examined breast cancer outcomes.

**Limitations**—Good quality evidence was sparse. Studies were small and confidence intervals were wide. Definitions of recall were absent or inconsistent.

**Conclusions**—Density ratings may be re-categorized on serial screening mammograms. Supplemental screening of women with dense breasts finds additional breast cancers, but increases false-positives. DBT may reduce recall rates. Supplemental screening impacts on breast cancer outcomes remain unclear.

**Primary Funding Source**—Agency for Healthcare Research and Quality

## INTRODUCTION

Dense breasts are defined by mammographic appearance. The American College of Radiology’s (ACR) Breast Imaging Reporting and Data System (BI-RADS), classifies breasts as almost entirely fatty (BI-RADS category a), scattered areas of fibroglandular density (category b), heterogeneously dense (category c), or extremely dense (category d).

About 27.6 million (43%) women aged 40 to 74 years in the United States have dense breasts; the majority of these are classified as category c (1). Higher breast density is

associated with decreased mammographic sensitivity and specificity and also with increased breast cancer risk. The relative hazard of breast cancer for women with dense breasts ranged from 1.50 (women aged 65–74 years) to 1.83 (women aged 40–49 years) in an analysis of 1,169,248 women enrolled in the Breast Cancer Surveillance Consortium (BCSC) (unpublished data). Increased breast density has been associated with hormone replacement therapy (HRT) use, younger age, and lower body mass index (2). Data on breast density and race-ethnicity are limited. In the United States, Asian women have higher breast density (3) but lower than average incidence of breast cancer (4). Increased breast density is not associated with higher breast cancer mortality among women with dense breasts diagnosed with breast cancer after adjusting for stage and mode of detection (5).

Supplemental breast cancer screening with additional screening modalities has been proposed to improve the early detection of breast cancers. Currently no clinical guidelines explicitly recommend use of supplemental breast cancer screening on women with dense breasts (6–9), but as of September 2015, 24 states have enacted legislation requiring women be notified of breast density with their mammography results; nine more states are considering mandatory notification (10) (Appendix Table 1, available at [www.annals.org](http://www.annals.org)). Most states require specific language distinguishing dense (BI-RADS c/d) from non-dense breasts and four states require that insurers cover subsequent examinations and tests for women with dense breasts (11–14). Federal legislation requiring breast density notification is pending (15).

This report summarizes a systematic review of current evidence on the reproducibility of BI-RADS breast density determinations, and on test performance characteristics and outcomes of supplemental screening of women with dense breasts with hand-held ultrasound (HHUS), automated whole breast ultrasound (ABUS), breast magnetic resonance imaging (MRI), and digital breast tomosynthesis (DBT). Mandatory reporting laws frame notification of women as dense/non-dense, so this review focused on this categorization.

## METHODS

The review protocol included an analytic framework with 4 key questions (Appendix Figure 1, available at [www.annals.org](http://www.annals.org)). Detailed methods, including search strategies, detailed inclusion criteria, and excluded studies are available in the full evidence report (16).

### Data Sources and Searches

MEDLINE, PubMed, Embase, and the Cochrane Library were searched for relevant English-language studies published between January 2000 and June 2015. We reviewed reference lists from retrieved articles and references suggested by experts.

### Study Selection

Two investigators independently reviewed abstracts and full-text articles for inclusion according to predetermined criteria (EPW and JHT for KQ1, JM and JJF for KQs 2–4). Included studies examining the reproducibility of BI-RADS breast density categorization focused on asymptomatic women aged 40 years or older undergoing digital or film mammography. Included studies on supplemental screening with HHUS, ABUS, MRI, or

DBT reported outcomes for asymptomatic women with dense breasts aged 40 years and older. In studies focused primarily on women at high risk for breast cancer, including those with pre-existing breast cancer or high-risk breast lesions (such as ductal carcinoma in situ [DCIS], atypical hyperplasia (ADH), lobular carcinoma in situ [LCIS]), BRCA mutation carriers, familial breast cancer syndromes, or prior chest-wall radiation, or including women with non-dense breasts, we analyzed the relevant subset when available in the publication or provided by the authors.

*A priori* inclusion criteria limited studies on BI-RADS reproducibility to fair- or good-quality randomized controlled trials (RCTs), cohort studies, or test sets involving multiple blind readings by at least three readers. Studies on test performance characteristics and outcomes of supplemental screening modalities were limited to fair- or good-quality RCTs, cohort studies, or diagnostic accuracy studies with reference standards applied to all subjects. We examined sensitivity, specificity, positive predictive values (PPV), negative predictive values (NPV) and available clinical outcomes including cancer detection rates, recall rates and biopsy rates. We defined recall as the need for any additional diagnostic testing after supplemental screening, including imaging and biopsy.

### Data Extraction and Quality Assessment

Two investigators (EPW and JHT for KQ1, JM and JJF for KQs 2–4) critically appraised all included studies independently using the USPSTF's design-specific criteria (17), supplemented with the National Institute for Health and Clinical Excellence methodology checklists (18) and the QAREL tool for assessing diagnostic reliability (19). According to USPSTF criteria, a good-quality study generally met all pre-specified criteria; fair-quality studies did not meet all criteria but had no important limitations. Poor quality studies had important limitations that could invalidate results (inadequate or biased application of reference standard; population limited to very high-risk patients).

### Data Synthesis and Analysis

When available or provided by the authors, we extracted results of supplemental screening for subgroups of women with dense breasts excluding those with other risk factors for breast cancer. We calculated the sensitivity and specificity of the supplemental breast screening tests for women with negative mammography. Only those cancers detected by the supplemental test after negative mammography and those found at interval follow-up were included. Hence the values reported represent the sensitivity and specificity for detection of additional cancers in women with negative mammography. Similarly, we defined cancer detection rates, recall rates and biopsy rates to include only those additional cancers, recalls, and biopsies related to supplemental screening after negative mammography. Meta-analysis was not performed due to few good quality studies.

### Role of Funding Source

This research was funded by AHRQ under a contract to support the work of the USPSTF. The investigators worked with USPSTF members to develop and refine the scope, analytic frameworks, and key questions. AHRQ had no role in study selection, quality assessment, synthesis, or development of conclusions. AHRQ provided project oversight; reviewed the

draft report; and distributed the draft for peer review, including to representatives of professional societies and federal agencies. AHRQ performed a final review of the manuscript to ensure that the analysis met methodological standards. The investigators are solely responsible for the content and the decision to submit the manuscript for publication.

## RESULTS

The literature search yielded 2,067 unique citations; 128 full text articles considered potentially relevant were reviewed to identify 24 unique studies meeting inclusion criteria (Appendix Figure 2, available at [www.annals.org](http://www.annals.org)). Table 1 provides the characteristics of included studies. No studies addressed the impact of supplemental screening (compared to women without supplemental screening) on breast cancer morbidity or mortality.

### BI-RADS Density Determination Accuracy and Reliability

Absent a gold standard for breast density, studies could not evaluate the accuracy of BIRADS density determinations. Five studies reported repeated assignment of categorical BI-RADS breast density classification by the same or different radiologists, altogether including >440,000 women, almost all with data from two sequential screening mammograms. To reflect current U.S. practice, we only included studies based on the BI-RADS density categories. The three largest studies were set in the United States, two using data from the Breast Cancer Screening Consortium (BCSC) (20, 22); the third presented findings from community radiologists conducting repeated readings of a large screening test set (24). Two other small studies (not discussed here) were based on mammographic screening programs in Spain (21) and Italy (23). All U.S.-based studies reflected community practice by use of clinical readings from community screening programs or test set readings by practicing community radiologists without additional training.

Overall group prevalence of BI-RADS density ratings were similar across initial and subsequent exams among community radiologists (Appendix Table 2, available at [www.annals.org](http://www.annals.org)), but there was greater disagreement at the individual level; on subsequent screening exams approximately one in five women (23%) were categorized into a different BI-RADS density category (a, b, c, d) by the same radiologist, while approximately one in three were categorized differently when the subsequent exam was read by a different radiologist (Table 2). Considering clinical interpretations that combine categories (“dense” representing those with BI-RADS c or d and “non-dense” representing BI-RADS a or b), 13–19% of women were reclassified into a different breast density category on their subsequent screening mammogram (Table 2).

These average estimates do not reflect greater extremes seen among outlier radiologists. Among 34 community radiologists reading sequential exams in the same women (22), readers assigned the same BI-RADS density assessment on both mammograms 77% of the time, on average; however, individual reader’s agreement between repeated ratings ranged from 62–87% (data not shown). In a study assessing repeat as well as cross-reader assignment of BI-RADS density categories by 19 radiologists in a test set of 341 exams, radiologists assigned the same BI-RADS density assignment 82% of the time, on average, although individual readers varied from 66–95% (24).

In community settings, 19–22% of exams initially classified as dense were subsequently reclassified as non-dense, while 10–16% of initially non-dense exams were reclassified as dense (Table 2). In contrast, initial clinical readings for a test set showed a higher percentage reclassified from non-dense to dense than vice versa. Across studies, the most commonly assigned breast density categories (b or c) were also those most likely to be reclassified on subsequent examination (Table 2), representing a clinical reclassification between non-dense and dense. Radiologists tended to agree with their own previous assessments of density better than with those made by other readers, although there was substantial variability among pairs of readers due to outliers (more details in full report (16)). These results are most applicable to post-menopausal women or those aged 50 years and older, since these women comprised between 71–100% of the study samples.

### Test Performance Characteristics of Supplemental Screening Technologies in Women with Dense Breasts

Nine studies reported test performance characteristics for supplemental screening with HHUS, ABUS and MRI among women with negative mammography (Table 2 and Appendix Figures 3 and 4, available at [www.annals.org](http://www.annals.org)). No studies reported test performance characteristics of DBT for women with dense breasts.

**Hand-Held Ultrasound (HHUS) and Automated Whole Breast Ultrasound (ABUS)**—Two good-quality studies (from the United States (25) and Italy (26)) and three fair-quality studies (29, 30, 34) reported on HHUS, and one fair-quality study from the United States (37) reported on ABUS (Table 3). We found no studies reporting variation in performance of these modalities by patient age and other breast cancer risk factors among women with dense breasts. Both good-quality studies applied consistent reference standards to identify interval cancers, and included over 1,000 women. The Italian study included women who self-referred to a charity-funded breast clinic, and reported findings separately by breast density category. The United States study included only women with dense breasts but many women also had additional major risk factors. The authors provided data for the subset of women without major risk factors. Additional details on all included studies are found in the full report (16).

Among women with dense breasts after recent negative screening mammography, the sensitivity of HHUS in the two good quality studies for detecting all breast cancer (including DCIS and invasive cancers) ranged from 0.80 (95% CI, 0.65 to 0.91) to 0.83 (95% CI, 0.59 to 0.96) (25, 26) and specificity ranged from 0.86 (95% CI, 0.85 to 0.88) to 0.95 (95% CI, 0.94 to 0.95) (25, 26). Sensitivity and specificity for invasive cancers was similar (25, 26). PPV in the good quality studies ranged from 0.03 to 0.08; NPV was 0.99 (25, 26). A single fair-quality study found that ABUS had similar performance characteristics to HHUS among women with dense breasts and negative mammography (37).

**MRI**—Three good-quality studies (25, 38, 39) reported test characteristics of supplemental MRI screening (Table 3). These studies included many women with elevated risk of breast cancer. In two studies, authors provided us with unpublished data for the study subgroup of women with dense breasts, excluding women at very high-risk due to BRCA 1/2 mutations,



chest irradiation, or personal histories of breast cancer (25, 39). In both, women had also had recent negative screening with HHUS. The third study included stratified results based on risk factors (38).

Among these subgroups of lower risk women with dense breasts, the sensitivity of MRI screening (after negative mammography) for all breast cancer ranged across studies from 0.75 (95% CI, 0.35 to 0.97) to 1.00 (95% CI, 0.59 to 1.00) (25, 38, 39). Specificity also varied, ranging from 0.78 (95% CI, 0.73 to 0.83) (25) to 0.93 (95% CI, 0.87 to 0.97) (39). PPV ranged from 0.03 to 0.33 and NPV's were 0.99 to 1.00.

### Cancer Detection and Recall Rates with Supplemental Screening

In general, supplemental screening following negative screening mammography consistently detected additional breast cancers, most of which were invasive. Eighteen studies reported rates of additional cancers detected, and most also reported recall and biopsy rates associated with supplemental screening (Table 4, Appendix Figure 5, available at [www.annals.org](http://www.annals.org)). With the possible exception of DBT, supplemental testing led to many additional recalls and biopsies.

**Hand-Held Ultrasound (HHUS) and Automated Whole Breast Ultrasound (ABUS)**—Seven studies reported HHUS cancer detection rates (27, 28, 31–33) and three studies reported on ABUS (35–37). The two good-quality studies of HHUS consistently estimated an all cancer detection rate after negative mammography of 4.4 per 1,000 exams (95% CI, 2.5 to 7.2) (25, 26), with invasive cancers comprising 93% (25) and 88% (26) of detected cancers. In the same women, mammography cancer detection rates were 4.7 per 1000 exams in the United States' study (25) and 2.8 per 1,000 exams in the Italian study (26). Only the United States' study reported the recall rate for supplemental HHUS: 14 % (95% CI, 12.7 to 15.1%) (25).

Three fair-quality studies reported cancer detection rates for ABUS. Cancer detection rates after negative mammography ranged from 1.9 to 15.2 per 1,000 exams (36, 37). In comparison, the cancer detection rate from mammography alone in one of these studies was 4.3 per 1000 exams (37). Recall rates varied between the studies from 2% (95% CI, 1.1 to 2.0%) to 14% (95% CI, 12.9 to 14.0%) (35, 36).

**MRI**—In three good-quality studies of MRI after negative mammography, breast cancer detection rates varied from 3.5 (95% CI, 1.3 to 7.6) to 28.6 per 1,000 exams (95% CI, 5.9 to 81.2) (25, 38, 39), with small numbers of cancers detected (range: 2–7). In comparison, mammography cancer detection rates in two of these studies for women with dense breasts were 4.1 and 7.0 per 1,000 exams (25, 38). Invasive breast cancers were 67% and 86% of detected cancers as reported by two studies (25, 39). Notably, women in these studies likely had higher breast cancer risk than the general population of women with dense breasts. In a good-quality United States study evaluating supplemental HHUS and MRI, among 334 women without BRCA mutations or prior breast cancer, after three screening rounds with negative mammography and HHUS over 24 months, screening breast MRI identified six additional invasive cancers (25).

Recall rates ranged from 9% (95% CI, 4.0 to 15.7%) to 23% (95% CI, 18.9 to 28.3%); the rate was highest in the study with three rounds of screening (25, 39). Biopsy rates were not reported separately for subgroups of women without increased risk. Because two of the studies reported on only one round of screening, the cumulative effect of recall for additional imaging and biopsy would likely increase with additional screening rounds.

**Digital Breast Tomosynthesis (DBT)**—Four fair-quality studies of DBT (three in the United States (41–43) and one in Italy (40)) reported on screening populations of women with dense breasts. All United States studies were single-site, retrospective studies, generally focused on outcomes before and after DBT introduction. In one study, breast cancer risk among women was described as above-average (41); other studies did not report on risk factors (40, 42, 43). Three studies reported cancer detection rates with digital mammography alone ranging from 4.0 to 5.2 per 1,000 exams (40, 42, 43). With DBT, combined detection ranged from 5.4 (95% CI, 3.5 to 7.9) to 6.9 (95% CI, 4.8 to 9.6) per 1,000 exams (42, 43). A single study reported 67% of cancers detected with combined DBT and mammography were invasive, the same proportion as with mammography alone (42). Recall rates with DBT in three retrospective United States studies ranged from 7% (95% CI, 6.2 to 7.7%) to 11% (95% CI, 10.0 to 11.7%), compared to 9% (95% CI, 8.4 to 11.0%) to 17% (95% CI, 15.0 to 18.2%) with digital mammography alone (41–43).

### Harms of Breast Density Notification

Only one study, a good-quality Canadian RCT, examined the effects of notification to women with normal screening results that their mammograms showed dense breasts (44). Women randomized to the intervention group (n=285) received a report of their breast density with letters summarizing their mammography results and a pamphlet on breast cancer risk factors including density. No supplemental screening was recommended. Women randomized to the control group (n=333) received notification of mammography results without information on breast density. At 4 weeks, more women in the intervention group had statistically significantly increased knowledge of breast density (25% in the intervention group vs 8% in the control group) and were more likely to perceive themselves as having elevated breast cancer risk. These differences did not persist at 6 months. There were no differences between groups in psychological distress, breast cancer worry, or preoccupation with breast cancer.

### Other Harms of Supplemental Screening

In studies of supplemental screening with HHUS and ABUS, over 90% of positive tests were false-positive, and in MRI studies 66–97% of all positive tests were false positives. Although no studies specifically addressed harms of supplemental screening in women with dense breasts, harms stemming from false-positive results are likely to be at least equivalent to those from mammography (45). We found no studies of whether focus on breast density distracts from assessment of other risk factors for breast cancer. Use of gadolinium contrast required for breast MRI has been associated with nephrogenic systemic fibrosis in patients with acute kidney injury or chronic kidney disease, but we found no reports of this adverse effect specifically related to breast MRI. The ACR recommends screening with serum creatinine prior to administration of gadolinium for those aged over 60 years and older with



hypertension, diabetes, or history of renal disease (46). Harms from DBT could come from additional breast radiation exposure (40–43, 47).

## DISCUSSION

We examined consistency of categorical BI-RADS breast density determinations in U.S. community practices since this is the system recommended by the ACR and written into most of the legislative mandates. Based on large, community practice-based studies, BI-RADS density assessments at a population level were generally consistent across sequential examinations by the same or different readers, but there was important variability among readings for individual women. Approximately 80% of examinations received a “b” or “c” BI-RADS density assessment; these categories were also most likely to be reassessed differently, whether on a separate reading of the same exam or on a subsequent examination, and whether read by the same or a different reader. As a result, across studies a sizeable 13–19% of women (13–19) were reclassified from “non-dense” to “dense” or vice versa. In these instances, mandated communications about elevated breast cancer risk or the need for additional clinical screenings could provide inconsistent information for the same woman in the span of 2 to 3 years.

Change in breast density findings can occur due to multiple factors stemming from the woman being examined, the qualitative nature of the technique, and radiologist variability in interpretation of the exams. The studies we examined tried to control for within-woman biological factors, suggesting that most of the variation in breast density assessment reflects within and between radiologist variability in density interpretation and the limitations of the current BI-RADS approach. Concerns about BI-RADS breast density determinations are a major impetus for research examining other methods for assigning breast density, including automated volumetric estimates, ultrasonographic assessments, and other computer-assisted methods. Although variability is reduced by use of double readings, which is widely practiced in Europe (40), this approach is impractical in the United States due to workforce requirements. The introduction of standards and quality measures related to breast density categorization could help to minimize potential harms associated with variable breast density categorizations.

When combined with mandated direct-to-consumer communications, variability in breast density assignments may lead to unintended consequences. Reclassification from one overall category to another (e.g., “dense” to “not-dense” or vice versa) may undermine a woman’s confidence in the screening process and leave her uncertain about her risk for breast cancer, while the opposite reclassification may alarm women unnecessarily or prompt supplemental screening tests of uncertain value. The ACR has publically expressed similar cautions about benefits, possible harms, and unintended consequences for the communication of breast density assessments to women (48).

Few studies evaluated test performance of supplemental screening tests for women with dense breasts. In the studies identified, the sensitivity of supplemental MRI screening after negative screening mammography appeared generally higher than HHUS screening, but even though we examined subsets of women without specific risk factors, we suspect that, in

general, these women were at higher risk. Studies of MRI were small and variable in their sensitivity estimates. No study directly compared sensitivity of supplemental screening modalities among women with dense breasts. Specificity of supplemental screening modalities was similar, and PPV was low. We identified only one study of ABUS and no studies of DBT test performance in women with dense breasts. No studies examined the effects of age or other breast cancer risk factors on supplemental test performance characteristics in women with dense breasts. No studies reported on breast cancer morbidity and mortality outcomes.

Evidence on harms of supplemental screening was also sparse. Added to digital mammography, DBT more than doubles the radiation exposure from each screening exam (49–51). New estimates of cancers induced by radiation from breast imaging have recently been reported (47). Technology that allows reconstruction of the 2-D breast images can reduce radiation exposure but is not widely disseminated (49). We found no reports of adverse effects from use of gadolinium contrast for breast MRI, but consideration should be given to a tracking mechanism for this potentially severe, though rare adverse effect. Potential harms resulting from overdiagnosis of breast cancer through supplemental screening can only be identified through rigorous prospective studies with long term follow-up.

### Limitations

Our review was limited to studies published in English; studies published in other languages may have met inclusion criteria, although applicability to U.S. practice could be limited. For applicability and feasibility concerns, we focused only on BI-RADS breast density assessment. Studies did not examine the underlying reasons for variability in BIRADS assessment within or between radiologists, nor did they evaluate any interventions to reduce the variability. The number, quality, and rigor of studies of diagnostic test characteristics and clinical outcomes were limited. Most studies lacked a complete reference standard, sufficient follow-up, or a clear description of follow-up, so diagnostic test performance characteristics could not be evaluated. Recall was often not clearly defined. No studies compared interval breast cancer rates, stage at diagnosis, or breast cancer mortality among two groups of women with dense breasts undergoing screening mammography with or without supplemental testing. No studies addressed the important potential risks of overdiagnosis and the associated harms of unnecessary treatment. Many studies included mixtures of women at increased breast cancer risk due to risk factors other than breast density, limiting the generalizability to the general screening population of women with dense breasts. Literature on ABUS and DBT for women with dense breasts was very limited, as was literature on the harms of breast density notification. Only one comparative study of cohorts with and without supplemental screening adjusted for differences between cohorts (42).

### Conclusions

Good-quality studies with United States radiologists show important reclassification between dense and non-dense breasts in women undergoing sequential screening examinations. Reclassification of breast density may introduce confusion or reduce confidence among women. Moving from a “dense” to a “non-dense” breast categorization

may result in different mandated communications in states with breast density notification, as well as fluctuation in clinical recommendations for supplemental screening.

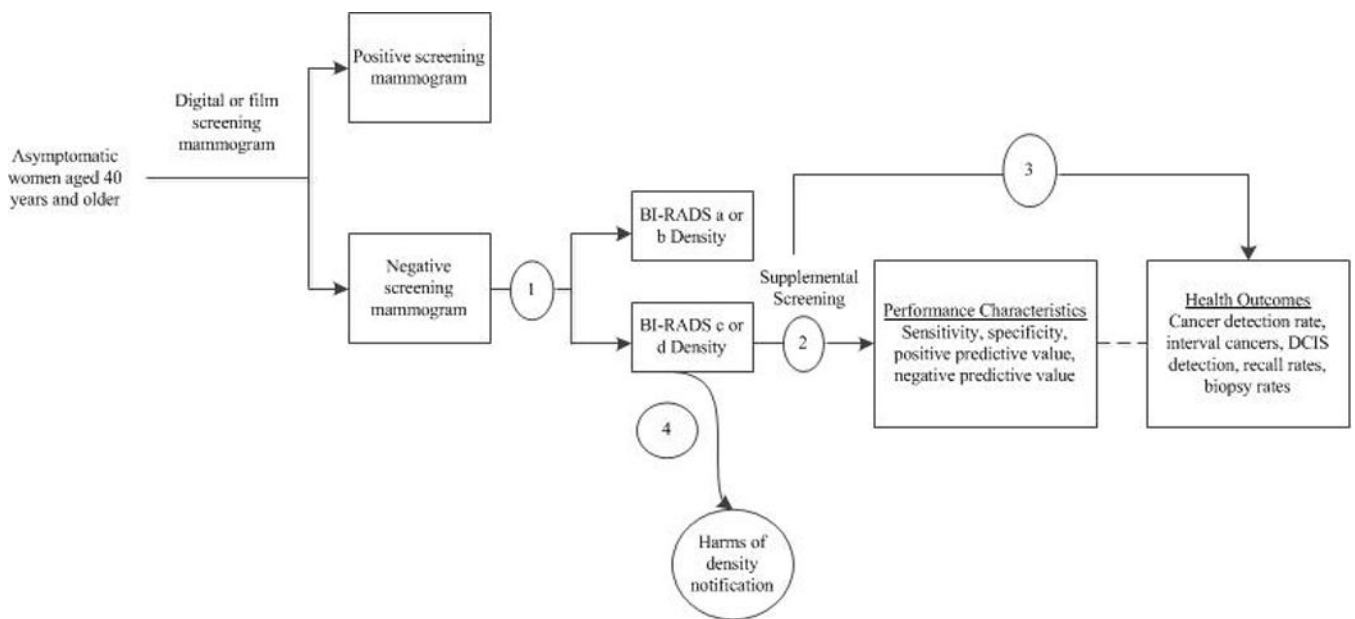
Limited evidence suggests that more breast cancers will be detected by supplemental HHUS and MRI screening of women with dense breasts, and most detected breast cancers will be invasive. Whether diagnosis of additional breast cancers by supplemental screening leads to improved clinical outcomes or what proportion of the cancers diagnosed represent overdiagnosis has not been evaluated. Supplemental testing of women with dense breasts with HHUS or MRI is associated with increased recall rates for diagnostic investigation among women without breast cancer. DBT may be associated with lower recall rates, but studies are few and retrospective. To define meaningful clinical outcomes of supplemental screening of women with dense breasts, well-designed, long-term prospective, comparative studies of supplemental screening are needed.

### Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: Agency for Healthcare Research and Quality staff, the U.S. Preventive Services Task Force, Joann Elmore, MD, MPH, Elizabeth Rafferty, MD, Jeffrey Tice, MD, Edward Sickles, MD, Barnett Kramer, MD, MH, Gretchen Gierach, PhD, and Gwendolyn Bryant-Smith, MD who provided expert and federal partner review of the report; Wendy Berg, MD, PhD and Christine Kuhl, MD for providing unpublished subgroup data; and Bruce Abbott, MLS and Guibo Xing, PhD at the University of California, Davis.

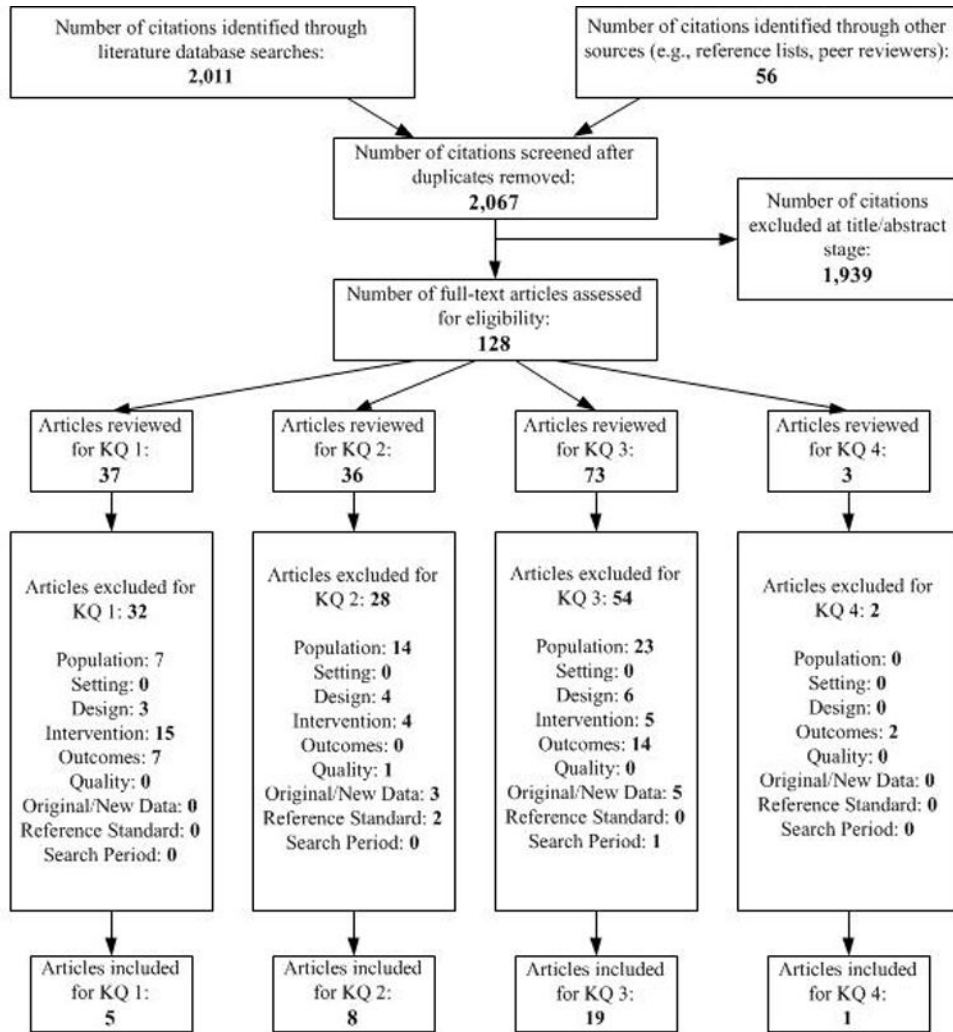
**Financial Support:** By AHRQ (contract HHSA-290-2012-00015I)

### Appendix



**Appendix Figure 1. Literature Flow Diagram**

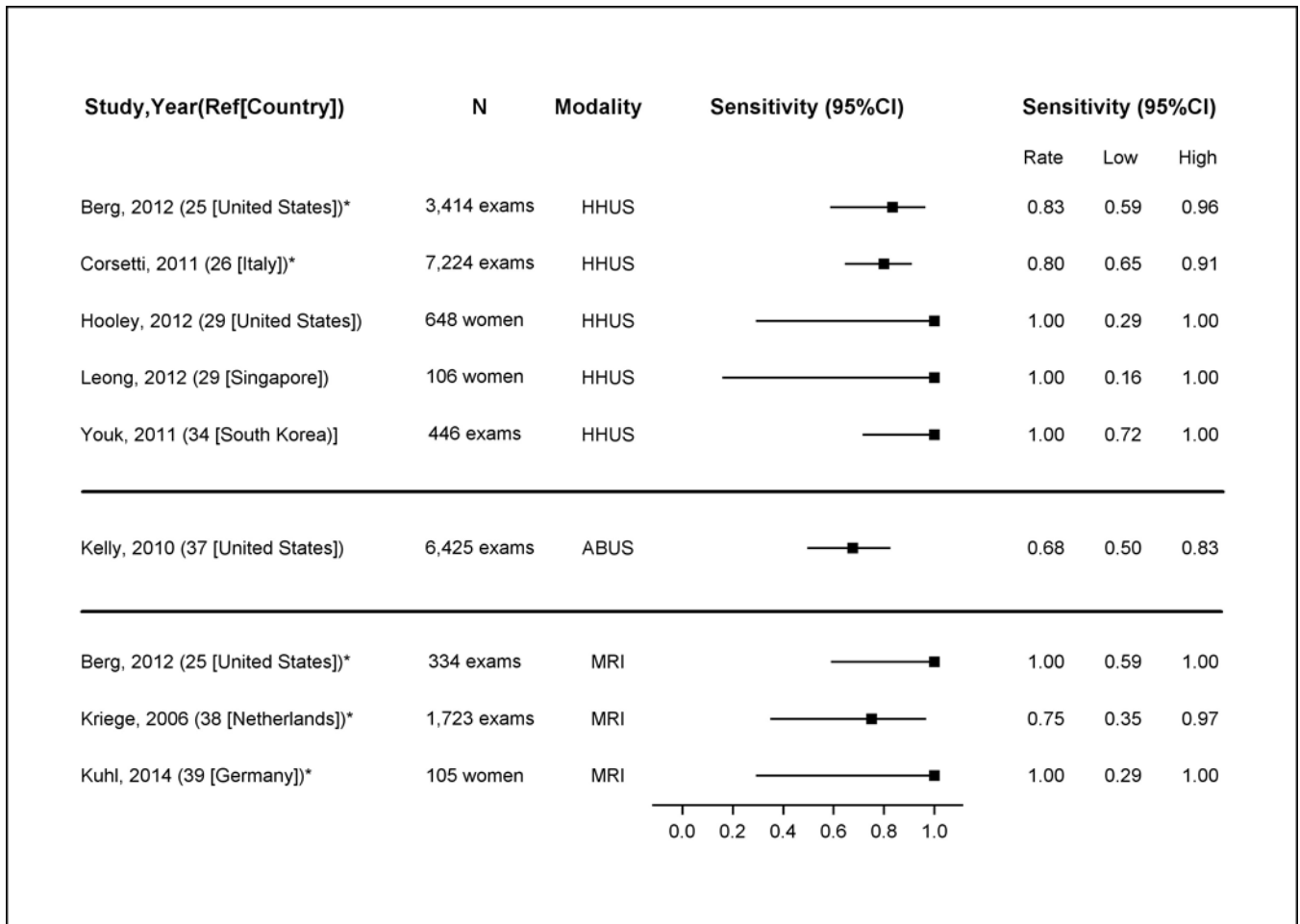
Abbreviations: BI-RADS=Breast Imaging Reporting and Data System; DCIS=ductal carcinoma in-situ; KQ=key question



- KQ1: What is the accuracy and reproducibility of BI-RADS determination of breast density?
- KQ2: What are the test performance characteristics of newer technologies for breast cancer screening when used as supplemental tests after a negative screening mammogram exam in women found to have dense breasts and how do these performance characteristics differ by age and risk factors?
- KQ3: When performed after a negative screening mammogram in women found to have dense breasts, what is the effectiveness of supplemental screening with breast ultrasound, MRI, or breast tomosynthesis on proximate clinical outcomes, including cancer detection rates, DCIS detection rates, stage at diagnosis, recall rates, biopsy rates, and interval cancer rates?
- KQ4: What are the harms associated with being identified as having dense breasts, including psychological and quality of life impacts and harms associated with supplemental screening evaluation, including evaluation of false positive results?

**Appendix Figure 2. Analytic Framework**

Abbreviations: BI-RADS=Breast Imaging Reporting and Data System; DCIS=ductal carcinoma in-situ; KQ=key question



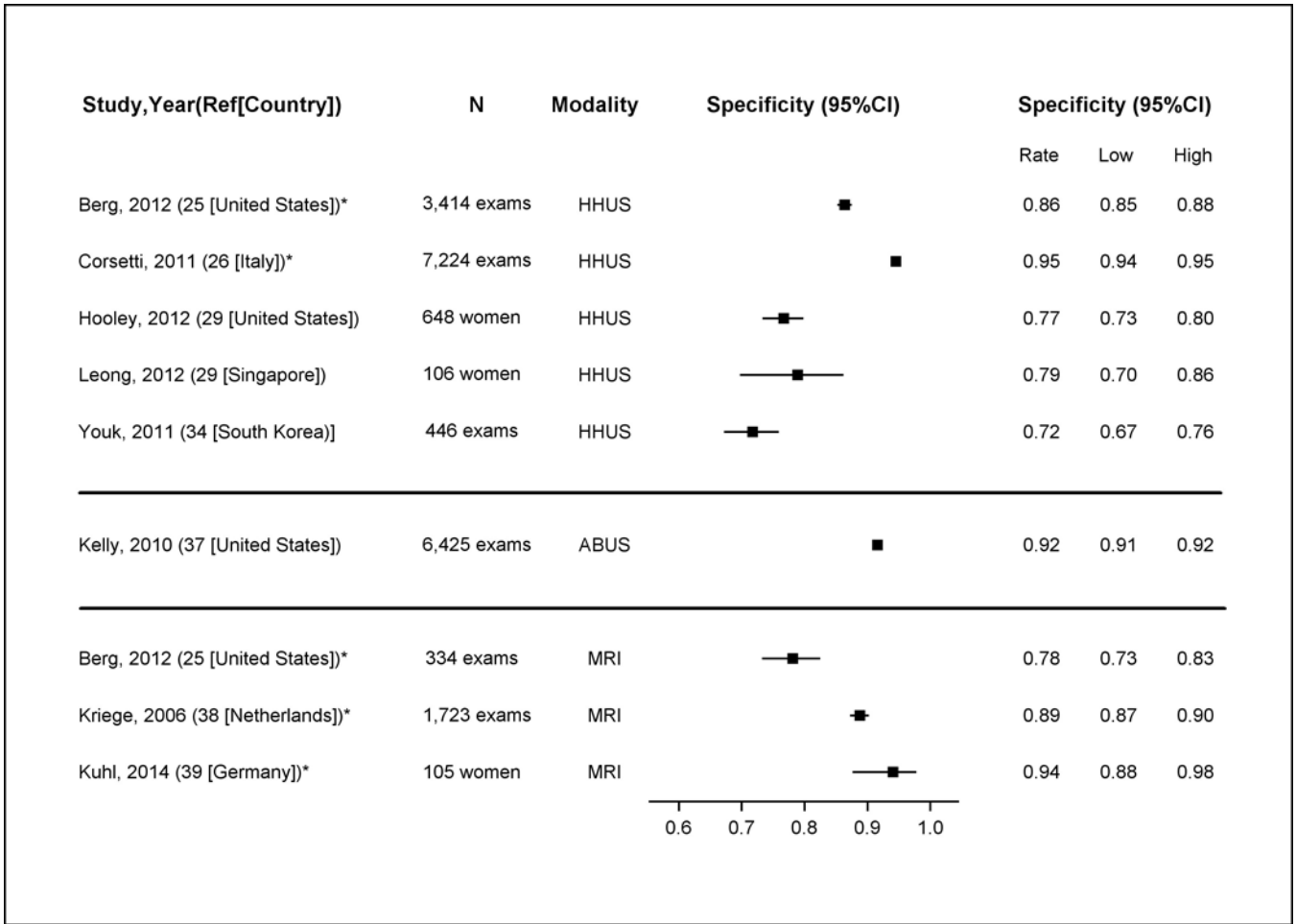
**Appendix Figure 3. Sensitivity of Supplemental Screening with Hand-Held Ultrasound (HHUS), Automated Whole Breast Ultrasound (ABUS) and MRI in Detecting Breast Cancer**

Note: these estimates include DCIS and invasive cancers

\*Good-quality study

Abbreviations: ABUS=automated whole breast ultrasound; CI=confidence interval;

HHUS=hand-held ultrasound



**Appendix Figure 4. Specificity of Supplemental Screening with Hand-Held Ultrasound (HHUS), Automated Whole Breast Ultrasound (ABUS) and MRI in Detecting Breast Cancer**

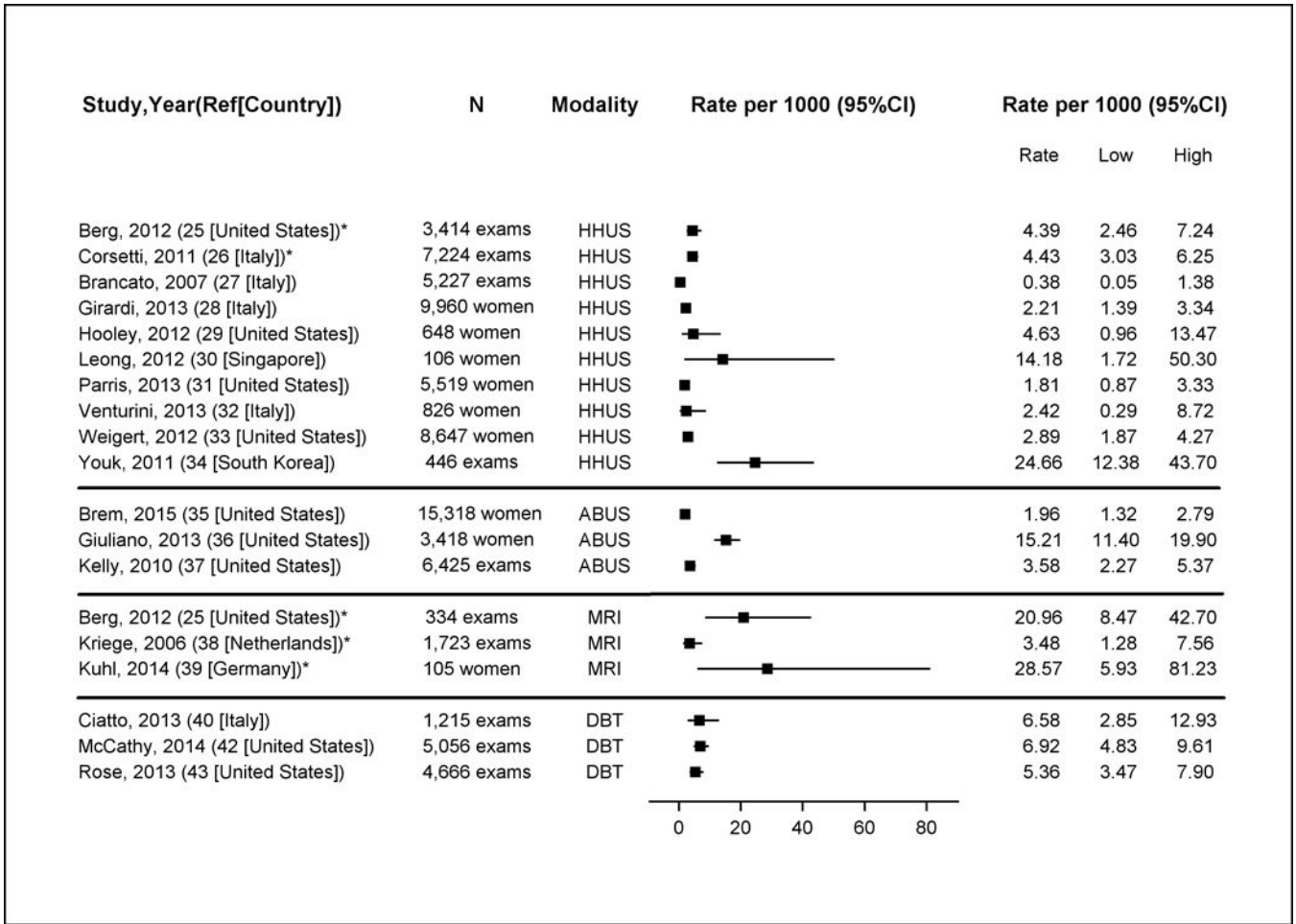
Note: these estimates include DCIS and invasive cancers

\*Good-quality study

Abbreviations: ABUS=automated whole breast ultrasound; CI=confidence interval;

HHUS=hand-held ultrasound





**Appendix Figure 5. Breast Cancer Detection Rates of Supplemental Screening with Hand-Held Ultrasound (HHUS), Automated Whole Breast Ultrasound (ABUS), MRI and digital breast tomosynthesis (DBT)**

Note: these estimates include DCIS and invasive cancers

\*Good-quality study

Abbreviations: ABUS=automated whole breast ultrasound; CI=confidence interval;

HHUS=hand-held ultrasound; DBT=digital breast tomosynthesis

**Appendix Table 1**

Breast Density Legislation in the United States

Status of Legislation*	Legislative Details	States
Pending Legislation	Drafting legislation mandating breast density notification	Florida, Maine, Illinois, Colorado, Vermont, Mississippi
	Introduced legislation mandating breast density notification <sup>7</sup>	Washington, Iowa, Indiana, Kentucky, South Carolina, Georgia
Enacted Legislation	Mandates patient notification about breast density	California, Arizona, Oregon, Nevada, Massachusetts, Minnesota, Texas, Alabama, Missouri, Tennessee, North Carolina, Virginia, Maryland, New Jersey, Pennsylvania, New York, Connecticut, Rhode Island, Hawaii,

Status of Legislation*	Legislative Details	States
		Michigan, Ohio, Louisiana, Delaware, North Dakota
	Requires specific language for patient notification	California, Arizona, Texas, Alabama, Missouri, Tennessee, North Carolina, Virginia, Maryland, New Jersey, Pennsylvania, New York, Connecticut, Rhode Island, Hawaii, Ohio, Michigan, Louisiana, Massachusetts
	Requires that <u>all</u> mammography reports provide information about breast density and the patient's current breast density level	Nevada, North Carolina, Maryland, Pennsylvania, Connecticut, Louisiana
	Requires that insurers cover appropriate medical examinations and tests for women with dense breasts	Illinois, Connecticut, New Jersey, Indiana

Source: (10)

\* As of September 2015

† During the 2015 legislative season

**Appendix Table 2**

**Consistency of BI-RADS Density Categories and Population Categorization**

Study Time between assessments	Study sample inclusion criteria	Assessments	a* (%)	b† (%)	c‡ (%)	d§ (%)
Harvey, 2013 (20) <36 mo	Women aged 40 years and older, with no history of breast cancer or reported use of hormone therapy at the time of exam or during the previous year, and had two or more screening mammographic examinations less than 36 months apart between January 1, 2000, and December 31, 2009 (n=435,751)	Exam 1	9.4	45.2	37.9	7.5
		Exam 2	10.2	45.1	37.2	7.2
Spayne, 2012 (24) 3–24 mo	Women who were postmenopausal <sup>//</sup> , with no history of breast cancer or reported use of hormone therapy, and had two or more film-screen screening or bilateral diagnostic mammograms including BI-RADS breast density assessments between January 1, 1996 and December 31, 2006 (n=11,755)	Exam 1	9.8	61.0	26.6	2.5
		Exam 2	9.2	60.2	28.1	2.5
Gard, 2015 (23) 6 mo	Women contributing examinations to the test set had a screening mammogram interpreted in the health care system between 1996 and 1998 and were enrolled in the system for at least 2 years following screening; the test set was designed to include about twice as many examinations of women with cancer as without, and roughly equal numbers of non-dense and dense examinations based on clinical interpretation (n=341)	Reading 1 <sup>¶</sup>	6.1	44.3	38.3	11.4
		Reading 2 <sup>#</sup>	4.5	39.2	47.0	9.3

\* Breast density category a = almost entirely fat

† Breast density category b = scattered fibroglandular densities

‡ Breast density category c = heterogeneously dense

§ Breast density category d = extremely dense

// Aged 55 years or older or reported having experienced natural menopause, having had both ovaries removed, or having more than 365 days elapse since their last menstrual period

¶ Clinical interpretation of exams prior to inclusion in test set

# Based on BI-RADS classification by the majority of readers for each exam in the test set

**Abbreviations:** BI-RADS=Breast Imaging-Reporting and Data System; mo=months

## References

1. Sprague BL, Gangnon RE, Burt V, Trentham-Dietz A, Hampton JM, Wellman RD, et al. Prevalence of mammographically dense breasts in the United States. *J Natl Cancer Inst.* 2014; 106(10)
2. Carmona-Sanchez E, Cuadros Lopez JL, Cuadros Celorrio AM, Perez-Roncero G, Gonzalez Ramirez AR, Fernandez Alonso AM. Assessment of mammographic density in postmenopausal women during long term hormone replacement therapy. *Gynecol Endocrinol.* 2013; 29(12):1067–70. [PubMed: 24004297]
3. Del Carmen MG, Hughes KS, Halpern E, Rafferty E, Kopans D, Parisky YR, et al. Racial differences in mammographic breast density. *Cancer.* 2003; 98(3):590–6. [PubMed: 12879477]
4. Centers for Disease Control and Prevention. Breast Cancer Rates by Race and Ethnicity 2014 [March 3, 2015]. Available from: <http://www.cdc.gov/cancer/breast/statistics/race.htm>
5. Gierach GL, Ichikawa L, Kerlikowske K, Brinton LA, Farhat GN, Vacek PM, et al. Relationship between mammographic density and breast cancer death in the Breast Cancer Surveillance Consortium. *J Natl Cancer Inst.* 2012; 104(16):1218–27. [PubMed: 22911616]
6. Lee CH, Dershaw DD, Kopans D, Evans P, Monsees B, Monticciolo 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. *J Am Coll Radiol.* 2010; 7(1):18–27. [PubMed: 20129267]
7. National Comprehensive Cancer Network (NCCN). Breast Cancer Screening and Diagnosis. 2014. [Accessed October 10, 2014]; Available from: [http://www.nccn.org/professionals/physician\\_gls/pdf/breast-screening.pdf](http://www.nccn.org/professionals/physician_gls/pdf/breast-screening.pdf)
8. American Cancer Society (ACS). American Cancer Society recommendations for early breast cancer detection in women without breast symptoms. 2014. [October 7, 2014]; Available from: <http://www.cancer.org/cancer/breastcancer/moreinformation/breastcancerearlydetection/breast-cancer-early-detection-ac-s-recs>
9. American College of Obstetricians and Gynecologists. ACOG Committee Opinion no. 593: Management of women with dense breasts diagnosed by mammography. *Obstet Gynecol.* 2014; 123(4):910–1. [PubMed: 24785636]
10. Are you Dense? Advocacy. D.E.N.S.E. ® State Efforts. 2015. [June 29, 2015]; Available from: <http://www.areyoudenseadvocacy.org/dense/>
11. Illinois General Assembly. (215 ILCS 5/356g) Sec. 356g. Mammograms; mastectomies. [October 10, 2014]; Available from: <http://www.ilga.gov/legislation/ilcs/documents/021500050k356g.htm>
12. 118th General Assembly of the State of Indiana. 2013. Senate Enrolled Act No. 414 [June 29, 2015]; Available from: <http://www.in.gov/legislative/bills/2013/PDF/SE/SE0414.1.pdf>
13. The State of Connecticut. An Act Requiring Communication of Mammographic Breast Density Information to Patients. 2009. Public Act No. 09.41 [Accessed October 10, 2014]; Available from: <http://www.cga.ct.gov/2009/ACT/PA/2009PA-00041-R00SB-00458-PA.htm>
14. New Jersey 215th Legislature. An Act Concerning Mammograms, amending P.L.1991, c.279 and P.L.2004, c.86, and supplementing Title 26 of the Revised Statutes and P.L.2003, c.193. 2014. [Accessed October 10, 2014]; Available from: [http://www.njleg.state.nj.us/2012/Bills/PL13/196\\_.PDF](http://www.njleg.state.nj.us/2012/Bills/PL13/196_.PDF)
15. Breast Density and Mammography Reporting Act of 2015. 114th Congress of the United States ed;
16. Melnikow, J.; Fenton, J.; Whitlock, E.; Miglioretti, DL.; Weyrich, M.; Thompson, J., et al. Adjunctive Screening for Breast Cancer in Women with Dense Breasts: A Systematic Review for the US Preventive Services Task Force. Rockville, MD: Agency for Healthcare Research and Quality; AHRQ Publication Number No. 14–05201-EF-2in press
17. U.S. Preventive Services Taskforce. US Preventive Services Taskforce Procedure Manual. Rockville, MD: Agency for Healthcare Research and Quality; 2013.
18. National Institute for Health and Clinical Excellence. The Guidelines Manual. London: National Institute for Health and Clinical Excellence; 2006.
19. Lucas NP, Macaskill P, Irwig L, Bogduk N. The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). *J Clin Epidemiol.* 2010; 63(8):854–61. [PubMed: 20056381]

20. Harvey JA, Gard CC, Miglioretti DL, Yankaskas BC, Kerlikowske K, Buist DS, et al. Reported mammographic density: film-screen versus digital acquisition. *Radiology*. 2013; 266(3):752–8. [PubMed: 23249570]
21. Redondo A, Comas M, Macia F, Ferrer F, Murta-Nascimento C, Maristany MT, et al. Inter- and intraradiologist variability in the BI-RADS assessment and breast density categories for screening mammograms. *British Journal of Radiology*. 2012; 85(1019):1465–70. [PubMed: 22993385]
22. Spayne MC, Gard CC, Skelly J, Miglioretti DL, Vacek PM, Geller BM. Reproducibility of BI-RADS breast density measures among community radiologists: a prospective cohort study. *Breast Journal*. 2012; 18(4):326–33. [PubMed: 22607064]
23. Bernardi G, Cavallaro G, Indinnimeo M, Fiore A, Basso L, D’Ermo G, et al. Usefulness of ultrasounds in the management of breast phyllodes tumors. *Il Giornale di chirurgia*. 2012; 33(3): 81–5. [PubMed: 22525552]
24. Gard CC, Bowles EJA, Miglioretti DL, Taplin SH, CM R. Misclassification of Breast Imaging Reporting and Data System (BI-RADS) mammographic density and implications for breast density reporting legislation. *Breast J*. 2015 epub ahead of print.
25. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, 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 – Journal of the American Medical Association*. 2012; 307(13):1394–404. [PubMed: 22474203]
26. Corsetti V, Houssami N, Ghirardi M, Ferrari A, Speziani M, Bellarosa S, 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*. 2011; 47(7):1021–6. [PubMed: 21211962]
27. Brancato B, Bonardi R, Catarzi S, Iacconi C, Risso G, Taschini R, et al. Negligible advantages and excess costs of routine addition of breast ultrasonography to mammography in dense breasts. *Tumori*. 2007; 93(6):562–6. [PubMed: 18338490]
28. Girardi V, Tonegutti M, Ciatto S, Bonetti F. Breast ultrasound in 22,131 asymptomatic women with negative mammography. *Breast*. 2013; 22(5):806–9. [PubMed: 23558244]
29. 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*. 2012; 265(1):59–69. [PubMed: 22723501]
30. 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*. 2012; 41(10):432–9.
31. Parris T, Wakefield D, Frimmer H. Real world performance of screening breast ultrasound following enactment of Connecticut Bill 458. *Breast Journal*. 2013; 19(1):64–70. [PubMed: 23240937]
32. Venturini E, Losio C, Panizza P, Rodighiero MG, Fedele I, Tacchini S, 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*. 2013; 268(2):347–55. [PubMed: 23579052]
33. Weigert J, Steenbergen S. The connecticut experiment: the role of ultrasound in the screening of women with dense breasts. *Breast Journal*. 2012; 18(6):517–22. [PubMed: 23009208]
34. Youk JH, Kim EK, Kim MJ, Kwak JY, Son EJ. Performance of hand-held whole-breast ultrasound based on BI-RADS in women with mammographically negative dense breast. *European Radiology*. 2011; 21(4):667–75. [PubMed: 20853108]
35. Brem RF, Tabar L, Duffy SW, Inciardi MF, Guingrich JA, Hashimoto BE, et al. Assessing improvement in detection of breast cancer with three-dimensional automated breast US in women with dense breast tissue: The sominsight study. *Radiology*. 2015; 274(3):663–73. [PubMed: 25329763]
36. Giuliano V, Giuliano C. Improved breast cancer detection in asymptomatic women using 3D-automated breast ultrasound in mammographically dense breasts. *Clinical Imaging*. 2013; 37(3): 480–6. [PubMed: 23116728]

37. Kelly KM, Dean J, Comulada WS, Lee SJ. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. *European Radiology*. 2010; 20(3):734–42. [PubMed: 19727744]
38. Kriege M, Brekelmans CT, Obdeijn IM, Boetes C, Zonderland HM, Muller SH, et al. Factors affecting sensitivity and specificity of screening mammography and MRI in women with an inherited risk for breast cancer. *Breast Cancer Research & Treatment*. 2006; 100(1):109–19. [PubMed: 16791481]
39. Kuhl CK, Schrading S, Strobel K, Schild HH, Hilgers RD, Bieling HB. Abbreviated Breast Magnetic Resonance Imaging (MRI): First Postcontrast Subtracted Images and Maximum-Intensity Projection-A Novel Approach to Breast Cancer Screening With MRI. *J Clin Oncol*. 2014; 32(22):2304–10. [PubMed: 24958821]
40. Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *The Lancet Oncology*. 2013; 14(7):583–9. [PubMed: 23623721]
41. 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*. 2013; 269(3):694–700. [PubMed: 23901124]
42. McCarthy AM, Kontos D, Synnestvedt M, Tan KS, Heitjan DF, Schnall M, et al. Screening outcomes following implementation of digital breast tomosynthesis in a general-population screening program. *J Natl Cancer Inst*. 2014; 106(11)
43. 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. American Journal of Roentgenology*. 2013; 200(6):1401–8. [PubMed: 23701081]
44. Bottorff JL, Ratner PA, Johnson JL, Hislop TG, Buxton JA, Zeisser C, et al. Women's responses to information on mammographic breast density. *Canadian Journal of Nursing Research*. 2007; 39(1): 38–57. [PubMed: 17450704]
45. Nelson, HD.; Cantor, A.; Humphrey, L.; Fu, R.; Pappas, M.; Daeges, M., et al. Screening for Breast Cancer: A Systematic Review to Update the 2009 U.S. Preventive Services Task Force Recommendation. Rockville, MD: Agency For Healthcare Research and Quality; 2015. AHRQ Publication Number 14-05201-EF-1
46. Besheli LD, Aran S, Shaqdan K, Kay J, Abujudeh H. Current status of nephrogenic systemic fibrosis. *Clin Radiol*. 2014; 69(7):661–8. [PubMed: 24582176]
47. Miglioretti, DL.; Lange, J.; van Ravesteyn, N.; van den Broek, JJ.; Lee, CI.; Melnikow, J., et al. Radiation-Induced Breast Cancer and Breast Cancer Death From Mammography Screening. Rockville, MD: Agency for Healthcare Research and Quality; AHRQ Publication No. 14-05201-EF-5 in press
48. American College of Radiology. ACR Statement on Reporting Breast Density in Mammography Reports and Patient Summaries April 24, 2012. [Accessed October 10, 2014]; Available from: <http://www.acr.org/About-Us/Media-Center/Position-Statements/Position-Statements-Folder/Statement-on-Reporting-Breast-Density-in-Mammography-Reports-and-Patient-Summaries>
49. Gur D, Zuley ML, Anello MI, Rathfon GY, Chough DM, Ganott MA, et al. Dose reduction in digital breast tomosynthesis (DBT) screening using synthetically reconstructed projection images: an observer performance study. *Academic Radiology*. 2012; 19(2):166–71. [PubMed: 22098941]
50. Feng SS, Sechopoulos I. Clinical digital breast tomosynthesis system: dosimetric characterization. *Radiology*. 2012; 263(1):35–42. [PubMed: 22332070]
51. Olgar T, Kahn T, Gosch D. Average glandular dose in digital mammography and breast tomosynthesis. *Rofo: Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin*. 2012; 184(10):911–8. [PubMed: 22711250]

**Table 1**

Characteristics of Included Studies

Study	USPSTF Quality Rating	Design	Country	Patients/Exams Analyzed	Follow-Up Period	Population Characteristics
<b>Breast Imaging-Reporting and Data System (BI-RADS) Density Assessment</b>						
Harvey, 2013 (20) Good		Cohort	United States	871,502 exams 435,751 women	<36 mo <sup>a</sup>	Age: 58.8y (mean)
Redondo, 2012 (21) Good		Stratified random sample	Spain	100 exams 100 women	6 mo <sup>a</sup>	Age: 50–64y (range)
Spayne, 2012 (22) Good		Cohort	United States	11,755 women	3–24 mo <sup>a</sup>	Age: 66y (median)
Bernardi, 2012 (23) Fair		Test set	Italy	100 exams 100 women	NA	Age: 43y (median)
Gard, 2015 (24) Fair		Test set	United States	341 women	6 mo <sup>a</sup>	NR
<b>Hand-Held Ultrasound (HHUS)</b>						
Berg, 2012 (25) Good		Test accuracy	United States	3,414 exams 1,216 women	12 mo	Age: 55.2y (mean) <sup>*</sup> Dense breasts: 100% Personal hx: 0% Family hx: NR
Corsetti, 2011 (26) Good		Test accuracy	Italy	7,224 exams 3,356 women	12 mo	Age: 55% <50y Dense breasts: 100% Personal hx: NR Family hx: NR
Brancato, 2007 (27) Fair		Cohort	Italy	5,227 women	NR	Age: 68% 40–49y Dense breasts: 100% Personal hx: NR Family hx: NR
Girardi, 2013 (28) Fair		Cohort	United States	9,960 women	12 mo	Age: 51.2y (mean) <sup>*</sup> Dense breasts: 100% Personal hx: 9.8% Family hx: NR
Hoohey, 2012 (29) Fair		Test accuracy	United States	648 women	15 mo	Age: 52y (mean) <sup>*</sup> Dense breasts: 100% Personal hx: NR Family hx: NR
Leong, 2012 (30) Fair		Test accuracy	Singapore	106 women	12–24 mo	Age: 45.1y (mean) Dense breasts: 100% Personal hx: 5% Family hx: 20.9%
Parris, 2013 (31) Fair		Cohort	United States	5,519 women	NR	Age: 53.6y (mean) Dense breasts: 89%



Study USPSTF Quality Rating	Design	Country	Patients/Exams Analyzed	Follow-Up Period	Population Characteristics
Venturini, 2013 (32) Fair	Cohort	Italy	826 women	NR	Personal hx: 6% Family hx: 42% Age: 100% <50y* Dense breasts: 100% Personal hx: NR Family hx: 24%*
Weigert, 2012 (33) Fair	Cohort	United States	8,647 exams 8,647 women	NR	Age: 54.4y (mean) Dense breasts: 100% Personal hx: NR Family hx: NR
Youk, 2011 (34) Fair	Test accuracy	South Korea	446 exams	24 mo	Age: 47.5y (mean)* Dense breasts: 100% Personal hx: 0% Family hx: NR
<b><u>Automated Whole Breast Ultrasound (ABUS)</u></b>					
Brem, 2015 (35) Fair	Cohort	United States	15,318	12 mo	Age: 53.3y (mean) Dense breasts: 100% Personal hx: 3.6% Family hx: 44.8%
Giuliano, 2013 (36) Fair	Cohort	United States	3,418 women	12 mo	Age: NR Dense breasts: 100% Personal hx: 0% Family hx: 0%
Kelly, 2010 (37) Fair	Test accuracy	United States	6,425 exams 4,419 women	12 mo	Age: 53y (mean)* Dense breasts: 68% Personal hx: 10%* Family hx: : 30%
<b><u>Magnetic Resonance Imaging (MRI)</u></b>					
Berg, 2012 (25) Good	Test accuracy	United States	334 exams	12 mo	Age: 56.8y (mean) Dense breasts: 100% Personal hx: NR Family hx: NR
Krieger, 2006 (38) Good	Test accuracy	Netherlands	1,723 exams	12 mo	Age: 40y (mean)* Dense breasts: 100% Personal hx: NR Family hx: 100%
Kuhl, 2014 (39) Good	Test accuracy	Germany	105 women	24 mo	Age: 53.2y (mean) Dense breasts: 100% Personal hx: NR Family hx: NR
<b><u>Digital Breast Tomosynthesis (DBT)</u></b>					
Ciatto, 2013 (40) Fair	Cohort	Italy	1,215 women	No follow-up except on biopsy results	Age: 58y (median)* Dense breasts: 100%

Study	USPSTF Quality Rating	Design	Country	Patients/Exams Analyzed	Follow-Up Period	Population Characteristics
Haas, 2013 (41) Fair		Cohort	United States	4,794 exams	No follow-up except on biopsy results	Personal hx: NR Family hx: NR Age: 33.3% <40y* Dense breasts: 100% Personal hx: 5.5%* Family hx: NR
McCarthy, 2014 (42) Fair		Cohort	United States	8,545 exams	No follow-up except on biopsy results	Age: 70% >50y Dense breasts: 100% Personal hx: NR Family hx: NR
Rose, 2013 (43) Fair		Cohort	United States	11,675 exams	No follow-up except on biopsy results	Age: 54.2y (mean)* Dense breasts: 100% Personal hx: NR Family hx: NR

<sup>a</sup>time between mammographic assessments

\*Data is reflective of the entire study population, not necessarily the subgroup with dense breasts

Abbreviations: Hx = history; mo=months; NA=not applicable; NR=not reported; y=years

**Table 2**  
Potential Misclassification of BI-RADS Density Categorization by Density Categories

Study	Repeat exam readers	% of women receiving a different breast density classification at 2 <sup>nd</sup> exam (4 categories)	% of women receiving an opposite breast density classification at 2 <sup>nd</sup> exam (2 categories)	% of dense exams reclassified as non-dense* (c or d) to (a or b)	% of dense exams reclassified as non-dense† (c) to (a or b)	% of dense exams reclassified as non-dense‡ (c) to (b)	% of non-dense exams reclassified as dense§ (a or b) to (c or d)	% of non-dense exams reclassified as dense¶ (b) to (c or d)	% of non-dense exams reclassified as dense   (b) to (c)
Harvey, 2013 (20)	Different community radiologists (n=703)	32	18.7	22	21	20	16	16	15
Spayne, 2012 (22)	Same community radiologists (n=34)	23	12.6	19	19	18	10	10	10
Gard, 2015 (24)	Same community radiologists (n=19)	29	16.9	10	10	10	23	23	23

\* Categorized as “heterogeneously dense” or “extremely dense” at first exam and “almost entirely fat” or “scattered fibroglandular densities” at second exam

† Categorized as “heterogeneously dense” at first exam and “almost entirely fat” or “scattered fibroglandular densities” at second exam

‡ Categorized as “heterogeneously dense” at first exam and “scattered fibroglandular densities” at second exam

§ Categorized as “almost entirely fat” or “scattered fibroglandular densities” at first exam and “heterogeneously dense” or “extremely dense” at second exam

¶ Categorized as “scattered fibroglandular densities” at first exam and “heterogeneously dense” or “extremely dense” at second exam

|| Categorized as “scattered fibroglandular densities” at first exam and “heterogeneously dense” at second exam

**Abbreviations:** BI-RADS=Breast Imaging-Reporting and Data System

**Table 3**

Test Performance Characteristics for Supplemental Hand-Held Ultrasound (HHUS), Automated Whole Breast Ultrasound (ABUS), and MRI

<b>Study Quality</b>	<b>Sensitivity (95% CI)</b>	<b>Specificity (95% CI)</b>	<b>PPV (95% CI)</b>	<b>NPV (95% CI)</b>
<b><i>HHUS</i></b>				
Berg, 2012 (25) Good	0.83 (0.59 to 0.96)	0.86 (0.85 to 0.88)	0.03 (0.02 to 0.05)	1.00 (1.00 to 1.00)
Corsetti, 2011 (26) Good	0.80 (0.65 to 0.91)	0.95 (0.94 to 0.95)	0.07 (0.05 to 0.10)	1.00 (1.00 to 1.00)
Hooley, 2012 (29) Fair	1.00 (0.29 to 1.00)	0.77 (0.73 to 0.80)	0.02 (0.01 to 0.06)	1.00 (0.99 to 1.00)
Leong, 2012 (30) Fair	1.00 (0.16 to 1.00)	0.79 (0.70 to 0.86)	0.08 (0.02 to 0.26)	1.00 (0.96 to 1.00)
Youk, 2011 (34) Fair	1.00 (0.72 to 1.00)	0.72 (0.67 to 0.76)	0.08 (0.05 to 0.14)	1.00 (0.99 to 1.00)
<b><i>ABUS</i></b>				
Kelly, 2010 (37) Fair	0.68 (0.50 to 0.83)	0.92 (0.91 to 0.92)	0.04 (0.03 to 0.06)	1.00 (1.00 to 1.00)
<b><i>MRI</i></b>				
Berg, 2012 (25) Good	1.00 (0.59 to 1.00)	0.78 (0.73 to 0.83)	0.09 (0.04 to 0.17)	1.00 (0.99 to 1.00)
Kriege, 2006 (38) Good	0.75 (0.35 to 0.97)	0.89 (0.87 to 0.90)	0.03 (0.01 to 0.06)	1.00 (1.00 to 1.00)
Kuhl, 2014 (39) Good	1.00 (0.29 to 1.00)	0.94 (0.88 to 0.98)	0.33 (0.12 to 0.66)	1.00 (0.96 to 1.00)

**Abbreviations:** CI=confidence interval; NPV=negative predictive value; PPV=Positive predictive value

**Table 4**

Breast Cancer Detection Outcomes for Adjunctive Hand-Held Ultrasound (HHUS), Automated Whole Breast Ultrasound (ABUS), MRI, and Digital Breast Tomosynthesis (DBT)

Study Quality	Cancers Detected	Cancer Detection Rate (95% CI)	Recall Rate (%) (95% CI)	Biopsy Rate (%) <sup>*</sup> (95% CI)
<b>HHUS</b>				
Berg, 2012 Good (25)	15/3,414 exams	4.4 per 1,000 exams (2.5 to 7.2)	14% (12.7 to 15.1)	NR
Corsetti, 2011 (26) Good	32/7,224 exams	4.4 per 1,000 exams (3.0 to 6.2)	NR	6% (5.4 to 6.5)
Brancato, 2007 (27) Fair	2/5,227 women	0.4 per 1,000 women (0 to 1.4)	2% (1.7 to 2.5)	1% (0.9 to 1.6)
Girardi, 2013 (28) Fair	22/9,960 women	2.2 per 1,000 women (1.4 to 3.3)	NR	NR
Hooley, 2012 (29) Fair	3/648 women	4.6 per 1,000 women (1.0 to 13.5)	24% (20.3 to 27.0)	7% (5.2 to 9.4)
Leong, 2012 (30) Fair	2/106 women	18.9 per 1,000 women (1.7 to 50.3)	17% <sup>‡</sup>	13% (7.4 to 21.2) <sup>‡</sup>
Parris, 2013 (31) Fair	10/5,519 women	1.8 per 1,000 women (0.9 to 3.3)	NR	3% (2.8 to 3.8)
Venturini, 2013 (32) Fair	2/826 women	2.4 per 1,000 women (0.3 to 8.7)	10% (7.5 to 11.7)	1% (0.6 to 2.2)
Weigert, 2012 (33) Fair	25/8,647 women	2.9 per 1,000 women (1.9 to 4.3)	14% (13.1 to 14.5)	5% (4.4 to 5.3)
Youk, 2011 (34) Fair	11/446 exams	24.7 per 1,000 exams (12.4 to 43.7)	14% (10.6 to 17.2)	11% (8.2 to 14.3)
<b>ABUS</b>				
Brem, 2015 (35) Fair	30/15,318 women	1.9 per 1,000 exams (1.3 to 2.8)	14% (12.9 to 14.0)	4% (3.4 to 4.0)
Giuliano, 2013 (36) Fair	DM+ABUS: 52/3418 women DM: 19/4076 women	DM+ABUS: 15.21 per 1,000 women (11.4 to 19.9) DM: 4.7 per 1,000 women (2.8 to 7.3)	2% (1.1 to 2.0)	NR
Kelly, 2010 (37) Fair	23/6425 exams	3.6 per 1,000 exams (2.3 to 5.4)	9% (8.0 to 9.4)	NR
<b>MRI</b>				
Berg, 2012 (25) Good	7/334 exams	21 per 1,000 exams (8.5 to 42.7)	23% (18.9 to 28.3)	NR
Kriege, 2006 (38) Good	6/1723 exams	3.5 per 1,000 exams (1.3 to 7.6)	12 % (10.0 to 13.1)	NR
Kuhl, 2014 (39) Good	3/105 women	28.6 per 1,000 women (5.9 to 81.2)	9% (4.0 to 15.7)	NR
<b>DBT</b>				
Ciatto, 2013 (40)	DBT+DM: 8/1215 exams DM: 5/1215 exams	DBT+DM: 6.6 per 1,000 exams (2.9 to 12.9) DM: 4.1 per 1,000 exams (1.3 to 9.6)	DBT+DM: 7% (5.2 to 8.1) DM only: 7% (5.8 to 8.8)	NR
Haas, 2013 (41)	NR	NR	DBT+DM: 10% (8.6 to 10.9) DM: 17% (15.0 to 18.2)	NR
McCarthy, 2014 (42) Fair	DBT+DM: 35/5056 exams DM: 18/3489 exams	DBT+DM: 6.9 per 1,000 exams (4.8 to 9.6) DM: 5.2 per 1,000 exams (3.1 to 8.1)	DBT+DM: 11% (10.0 to 11.7)	NR

Study Quality	Cancers Detected	Cancer Detection Rate (95% CI)	Recall Rate (%) (95% CI)	Biopsy Rate (%) <sup>*</sup> (95% CI)
Rose, 2013 (43) Fair	DBT+DM: 25/4666 exams DM: 28/7009 exams	DBT+DM: 5.4 per 1,000 exams (3.5 to 7.9) DM: 4.0 per 1,000 exams (2.7 to 5.8)	DM: 13% (11.7 to 14.0) DBT+DM: 7% (6.2 to 7.7) DM: 9% (8.4 to 11.0)	NR

\* Biopsy rate includes needle aspiration, core needle, and open biopsies

<sup>†</sup>Data are based on the 106 women with complete follow-up (out of 141 total)

**Abbreviations:** CI=confidence interval; DM=digital mammography; NR=not reported

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript