



Online article and related content
current as of May 13, 2008.

Combined Screening With Ultrasound and Mammography vs Mammography Alone in Women at Elevated Risk of Breast Cancer

Wendie A. Berg; Jeffrey D. Blume; Jean B. Cormack; et al.

JAMA. 2008;299(18):2151-2163 (doi:10.1001/jama.299.18.2151)

<http://jama.ama-assn.org/cgi/content/full/299/18/2151>

Supplementary material	JAMA Report Video http://jama.ama-assn.org/cgi/content/full/299/18/2151/DC1
Correction	Contact me if this article is corrected.
Citations	This article has been cited 1 time. Contact me when this article is cited.
Topic collections	Radiologic Imaging; Women's Health; Women's Health, Other; Diagnosis; Oncology; Breast Cancer; Mammography; Ultrasonography Contact me when new articles are published in these topic areas.
Related Articles published in the same issue	The "Coming of Age" of Nonmammographic Screening for Breast Cancer Christiane K. Kuhl. <i>JAMA</i>. 2008;299(18):2203.

Subscribe
<http://jama.com/subscribe>

Permissions
permissions@ama-assn.org
<http://pubs.ama-assn.org/misc/permissions.dtl>

Email Alerts
<http://jamaarchives.com/alerts>

Reprints/E-prints
reprints@ama-assn.org

Combined Screening With Ultrasound and Mammography vs Mammography Alone in Women at Elevated Risk of Breast Cancer

Wendie A. Berg, MD, PhD

Jeffrey D. Blume, PhD

Jean B. Cormack, PhD

Ellen B. Mendelson, MD

Daniel Lehrer, MD

Marcela Böhm-Vélez, MD

Etta D. Pisano, MD

Roberta A. Jong, MD

W. Phil Evans, MD

Marilyn J. Morton, DO

Mary C. Mahoney, MD

Linda Hovanessian Larsen, MD

Richard G. Barr, MD, PhD

Dione M. Farria, MD, MPH

Helga S. Marques, MS

Karan Boparai, RT

for the ACRIN 6666 Investigators

EARLY DETECTION REDUCES DEATHS due to breast cancer. The US Preventive Services Task Force analysis of 7 randomized trials of mammographic screening found that the point estimate of the reduction in mortality from screening mammography was 22% in women aged 50 years or older and 15% among women between 40 and 49 years,¹ with some individual trials showing far greater benefits in both age groups and with any specific age distinction arbitrary. The magnitude of reduction in mortality seen in individual trials parallels reductions in size distribution² and rates of node-positive breast cancer.³

Mammography can depict calcifications due to malignancy, including ductal carcinoma in situ (DCIS). Invasive

For editorial comment see p 2203.

Context Screening ultrasound may depict small, node-negative breast cancers not seen on mammography.

Objective To compare the diagnostic *yield*, defined as the proportion of women with positive screen test results and positive reference standard, and performance of screening with ultrasound plus mammography vs mammography alone in women at elevated risk of breast cancer.

Design, Setting, and Participants From April 2004 to February 2006, 2809 women, with at least heterogeneously dense breast tissue in at least 1 quadrant, were recruited from 21 sites to undergo mammographic and physician-performed ultrasonographic examinations in randomized order by a radiologist masked to the other examination results. *Reference standard* was defined as a combination of pathology and 12-month follow-up and was available for 2637 (96.8%) of the 2725 eligible participants.

Main Outcome Measures Diagnostic yield, sensitivity, specificity, and diagnostic accuracy (assessed by the area under the receiver operating characteristic curve) of combined mammography plus ultrasound vs mammography alone and the positive predictive value of biopsy recommendations for mammography plus ultrasound vs mammography alone.

Results Forty participants (41 breasts) were diagnosed with cancer: 8 suspicious on both ultrasound and mammography, 12 on ultrasound alone, 12 on mammography alone, and 8 participants (9 breasts) on neither. The diagnostic yield for mammography was 7.6 per 1000 women screened (20 of 2637) and increased to 11.8 per 1000 (31 of 2637) for combined mammography plus ultrasound; the supplemental yield was 4.2 per 1000 women screened (95% confidence interval [CI], 1.1-7.2 per 1000; $P = .003$ that supplemental yield is 0). The diagnostic accuracy for mammography was 0.78 (95% CI, 0.67-0.87) and increased to 0.91 (95% CI, 0.84-0.96) for mammography plus ultrasound ($P = .003$ that difference is 0). Of 12 supplemental cancers detected by ultrasound alone, 11 (92%) were invasive with a median size of 10 mm (range, 5-40 mm; mean [SE], 12.6 [3.0] mm) and 8 of the 9 lesions (89%) reported had negative nodes. The positive predictive value of biopsy recommendation after full diagnostic workup was 19 of 84 for mammography (22.6%; 95% CI, 14.2%-33%), 21 of 235 for ultrasound (8.9%, 95% CI, 5.6%-13.3%), and 31 of 276 for combined mammography plus ultrasound (11.2%; 95% CI, 7.8%-15.6%).

Conclusions Adding a single screening ultrasound to mammography will yield an additional 1.1 to 7.2 cancers per 1000 high-risk women, but it will also substantially increase the number of false positives.

Trial Registration clinicaltrials.gov Identifier: NCT00072501

JAMA. 2008;299(18):2151-2163

www.jama.com

cancers, which can spread to lymph nodes and cause systemic metastases, are most often manifest as noncalcified masses⁴ and can be mammographically subtle or occult, particularly when the parenchyma is dense. Dense breast tissue is common. More than half of

women younger than 50 years⁵ have either *heterogeneously dense*, visually estimated as 51% to 75% glandular,⁶ or

Author Affiliations are listed at the end of this article.
Corresponding Author: Wendie A. Berg, MD, PhD, 10755 Falls Rd, Suite 440, Lutherville, MD 21093 (wendieberg@gmail.com).

extremely dense, visually estimated as more than 75% glandular⁶ breasts, as do at least one-third of women older than 50 years.⁵ In women with dense breasts, mammographic sensitivity may be as low as 30% to 48%,^{7,8} with much higher interval cancer rates^{7,9} and worse prognosis for resulting clinically detected cancers. Furthermore, dense breast tissue is itself a marker of increased risk of breast cancer on the order of 4- to 6-fold.¹⁰ In dense breasts, digital mammography has improved performance, with sensitivity increasing from 55% with screen film to 70% with digital in 1 large series using mammographic and clinical follow-up as a gold standard.¹¹ Digital mammography does not, however, eliminate the fundamental limitation that noncalcified breast cancers are often obscured by surrounding and overlying dense parenchyma.

In women younger than 50 years, the reduced benefit of mammographic screening is attributed to increased breast density, biologically more aggressive cancers, and reduced prevalence of disease. Using a screening interval of 12 months, rather than 24 months, should improve results with rapidly growing malignancies, even though dense tissue remains a major limitation to improving outcomes.¹² Methods to address improving detection despite dense breast tissue are needed.

Supplemental screening ultrasound has the potential of depicting small, node-negative breast cancers not seen on mammography,^{8,13-17} and its performance is improved in dense parenchyma.⁸ It is natural to expect that methods that improve the detection of small, node-negative cancers would further reduce mortality when performed in addition to screening mammography. However, direct evidence of a mortality reduction due to screening can only be generated in a large prospective randomized screening trial with mortality as an end point. Such trials are costly, require extensive infrastructure and resources, and are not practical under all contexts. Surrogate aims and end points, such as the diagnostic performance for the screening mo-

dality or the size and stage of breast cancers depicted, have been correlated with mortality outcomes,^{18,19} and can be used to project the mortality reduction if the screening modality were implemented.

Across 42 838 examinations from the 6 published single-center studies of screening ultrasound to date,^{8,13-17} 126 women (0.29%) were shown to have 150 cancers identified only on supplemental ultrasound.²⁰ Of 141 invasive cancers detected only on ultrasound, 99 (70%) were 1 cm or smaller in size.²⁰ In studies for which staging was detailed, 36 of 40 cancers (90%) depicted by ultrasonography alone were categorized as stage 0 or I.²⁰

Concerns remain, however, over the generalizability of such favorable results with screening ultrasound. In particular, there is concern for the operator dependence of freehand screening breast ultrasound because an abnormality must be perceived while scanning for it to be documented. Importantly, recent reports have shown that consistent breast ultrasound examination performance and interpretation is possible with minimal training.^{21,22} Other limitations to implementing widespread screening ultrasound include a shortage of qualified personnel to perform and interpret the examination and lack of standardized scanning protocols. These concerns have hampered use of screening ultrasound; 35% of surveyed facilities specializing in breast imaging offered it in 2005,²³ even though most facilities offering screening ultrasound will do so only on a limited basis.

In this study, we report a prospective, multicenter trial, randomized to sequence of performance of mammography and ultrasound, designed to investigate and validate the performance of screening ultrasound in conjunction with mammography, using a standardized protocol and interpretive criteria. This trial was designed to compare the diagnostic yield of screening breast mammography plus ultrasound with mammography alone in women at increased risk of breast cancer. Since beginning this trial, a multicenter study was

published from Italy in which 6449 women with dense breasts and negative mammogram results underwent screening ultrasound, with 29 cancers depicted by ultrasound (cancer detection rate, 0.45%).²⁴ The American College of Radiology Imaging Network (ACRIN) 6666 is the largest trial of screening ultrasound in which mammography and ultrasound have been performed and read independently, allowing detailed analysis of the performance of each modality separately and in combination and reducing potential biases in patient recruitment and interpretation of both mammography and ultrasound. Furthermore, we used standardized scanning and interpretive criteria (<http://www.acrin.org/TabID/153/Default.aspx>), which should facilitate generalizability of our results.

Unlike previous reports evaluating screening ultrasound, we chose to study a population at elevated risk of breast cancer. Supplemental screening in addition to mammography may be more cost-effective in such populations because the expected prevalence of disease is higher than it is for populations with no risk factors. Furthermore, patients at higher risk may be encouraged to begin screening at an earlier age when the tissue is denser and mammography is more limited in its benefits. Indeed, annual magnetic resonance imaging (MRI) is now recommended in addition to mammography for women at very high risk of breast cancer,²⁵ but it remains limited by high cost, required injection of contrast, reduced patient tolerance, and limited availability and expertise. Ultrasound is relatively inexpensive, requires no contrast, is well tolerated, and is widely available.

METHODS

Study Design

Participants were women at elevated risk of breast cancer (TABLE 1) who presented for routine annual mammography and provided written informed consent. Each participant underwent mammographic and ultrasonographic screening examinations in randomized

order with the interpreting radiologist for each examination masked to results of the other. Random assignment of screening order was stratified by site and block randomization with alternating block sizes of 6 and 8 used within each site. If the recommendation from the study mammography or ultrasound was for other than routine annual screening, an integrated mammography plus ultrasound interpretation was recorded by a qualified site investigator radiologist. Otherwise, if both ultrasound and mammography were interpreted as negative or benign, no separate integrated interpretation was performed, and the combination of mammography plus ultrasound was assumed to be negative. Management was based on recommendations from the integrated examination. If needed, targeted ultrasonographic or additional mammographic views were then performed and results, assessments, and recommendations were separately recorded. Results of repeat screening at 12 and 24 months after study entry are still being collected. Race and ethnic group were self-assigned from a list of options for ethnicity and a series of yes or no questions for race.

Web-based data capture and quality monitoring was conducted by ACRIN's Biostatistics and Data Management Center. For all analyses in this study, data were cleaned and locked as of May 14, 2007. The study received institutional review board approval from all participating sites; ACRIN and National Cancer Institute-Cancer Imaging Program approval; and data and safety monitoring committee review every 6 months.

Participant Population

A total of 2809 women were recruited from 21 sites between April 2004 and February 2006, of whom 2725 were eligible (FIGURE 1, Table 1). Women aged at least 25 years who presented for routine annual mammography were eligible to participate if they met uniform definitions of elevated risk (Table 1) as determined by study personnel and had heterogeneously dense or extremely dense parenchyma⁶ in at least 1 quadrant, either by prior mam-

mography report or by review of prior mammograms. Otherwise eligible women with no prior mammography were allowed to enroll under the rationale that such women would be high-risk young women presenting for baseline screening who would usually have dense breasts. Women were excluded if they had signs or symptoms of breast cancer, recent surgical or percutaneous image-guided breast interventional procedures or MRI or tomosynthesis of the breast(s) within the prior

12 months, or mammography or whole breast ultrasound fewer than 11 months earlier. Also excluded were women with breast implants and those who were pregnant, lactating, or planning to become pregnant within 2 years of study entry or who had known metastatic disease. We did not exclude women with prior breast cancer or basal or squamous cell skin cancer or in situ cervical cancer. Women with other prior cancers were eligible to enroll if they had been disease-free for at least 5 years.

Table 1. Participant Characteristics

	No. (%) of Participants ^a	
	Eligible (n = 2725)	Analysis Set (n = 2637)
Age at enrollment, y ^b		
Median (range)	55 (25-91)	55 (25-91)
Mean (SE)	55.14 (0.19)	55.23 (0.20)
Race/ethnicity ^c		
White	2519 (92.44)	2448 (92.83)
Hispanic or Latina ^d	274 (10.06)	259 (9.82)
Black or African American	100 (3.67)	90 (3.41)
Native Hawaiian or other Pacific Islander	4 (0.15)	4 (0.15)
Asian	95 (3.49)	88 (3.34)
American Indian or Alaskan Native	4 (0.15)	4 (0.15)
Unknown	11 (0.40)	11 (0.42)
Menopausal status ^e		
Premenopausal	629 (23.08)	601 (22.79)
Perimenopausal	188 (6.90)	179 (6.79)
Postmenopausal	1387 (50.90)	1353 (51.31)
Surgical menopause	492 (18.06)	482 (18.28)
Unknown/I cannot remember	22 (0.81)	22 (0.83)
Data missing	7 (0.26)	0 (0.0)
Risk factors ^f		
Personal history of breast cancer	1443 (52.95)	1400 (53.09)
Lifetime risk $\geq 25\%$ by Gail or Claus model	517 (18.97)	497 (18.85)
5-y Gail model risk $\geq 2.5\%$	411 (15.08)	403 (15.28)
$\geq 1.7\%$ and extremely dense breasts	230 (8.44)	223 (8.46)
ADH, ALH, LCIS, or atypical papilloma	84 (3.08)	83 (3.15)
Mutation in <i>BRCA1</i> or <i>BRCA2</i> genes ^g	24 (0.88)	23 (0.87)
Chest, mediastinal, or axillary irradiation ^e	8 (0.29)	8 (0.30)
Basis of eligibility unknown	8 (0.29)	0
Imaging history		
Mammogram before study entry, mo		
< 14	1977 (72.55)	1929 (73.15)
14-24	587 (21.54)	561 (21.27)
> 24	107 (3.93)	104 (3.94)
None	54 (1.98)	43 (1.63)
Prior breast ultrasound		
Targeted	1891 (69.40)	1834 (69.55)
Screening	304 (11.16)	293 (11.11)
Contrast-enhanced breast MRI	196 (7.19)	191 (7.24)

(continued)

Table 1. Participant Characteristics (cont)

	No. (%) of Participants ^a	
	Eligible (n = 2725)	Analysis Set (n = 2637)
Hormone use		
Current	148 (5.43)	143 (5.42)
Any prior	792 (29.06)	780 (29.58)
Never	1785 (65.50)	1714 (65.00)
Current chemoprevention at study entry		
Tamoxifen	338 (12.40)	329 (12.48)
Raloxifene	95 (3.49)	94 (3.56)
Aromatase inhibitor	142 (5.21)	141 (5.35)
Brassiere cup size		
A	449 (16.48)	439 (16.65)
B	983 (36.07)	949 (35.99)
C	856 (31.41)	830 (31.48)
D	312 (11.45)	304 (11.53)
DD	91 (3.34)	89 (3.38)
Other, specify	27 (0.99)	26 (0.99)
Data missing	7 (0.26)	0

Abbreviations: ADH, atypical ductal hyperplasia; ALH, atypical lobular hyperplasia; LCIS, lobular carcinoma in situ.

^aPercentages may not sum to 100 due to rounding.

^bThe age distribution for the eligible set was 43 (10th percentile), 48 (25th percentile), 55 (50th percentile), 62 (75th percentile), and 68 (90th percentile) years. The age distribution for the analysis set was 43 (10th percentile), 49 (25th percentile), 55 (50th percentile), 62 (75th percentile), and 68 (90th percentile) years.

^cParticipants may fall into more than 1 category. Ethnicity was self-selected from a list of options and race from a series of yes or no questions.

^dOf the Hispanic or Latina participants in the eligible and analysis sets, 80% were accrued at CERIM, Buenos Aires, Argentina. None of the black or African American participants in either set were accrued at CERIM.

^ePremenopausal women had their last menstrual period less than 1 month before registration. Perimenopausal women had their last menstrual period at least 1 month but less than 12 months before registration. Postmenopausal women had their last menstrual period more than 1 year before registration.

^fParticipants are listed only once for the risk factor that determined eligibility. *BRCA*-mutation carrier status and prior chest/mediastinal radiation were prioritized over personal history of breast cancer, which was prioritized over prior atypical biopsy result, then risk by Gail or Claus models (lifetime risk prioritized over 5 y risks >2.5% then 1.7% and extremely dense parenchyma, respectively).

^gNine women with *BRCA1* or *BRCA2* mutations also had a personal history of breast cancer, as did 4 with prior chest, or mediastinal radiation, or both.

Screening Methods

At least 2-view mammography was performed using either screen-film or digital mammography. Visually estimated overall mammographic breast density on study mammograms was recorded as less than 25%; 26% to 40%; 41% to 60%; 61% to 80%; or more than 80% dense. Computer-assisted detection was not permitted. Radiologist investigators who had successfully completed both phantom scanning²⁶ and mammographic and ultrasonographic interpretive skills tasks²⁷ performed separate, masked interpretations of mammographic and ultrasonographic examinations. Survey ultrasound was performed using high-resolution linear array, broad bandwidth transducers with maximum frequency of at least 12 MHz, with scanning in transverse and sagittal planes. Lesions other than simple cysts were imaged with and without spa-

tial compounding and power or color Doppler in orthogonal planes (typically radial and antiradial orientations). An image (with embedded clock time) was recorded on entering the ultrasound suite, at the beginning and end of ultrasonographic screening, and on leaving the suite to determine the time to scan and the total physician time in the room. Electively, the axilla could be scanned, and its inclusion was recorded. Investigators recorded ultrasonographic background echotexture and lesion features using Breast Imaging and Reporting Data System (BI-RADS): Ultrasound descriptors²⁸ and average breast thickness to the nearest centimeter.

Assessments for each lesion and for each breast overall were recorded on the expanded 7-point BI-RADS⁶ scale: 1, negative; 2, benign; 3, probably benign; 4a, low suspicion; 4b, intermediate sus-

picion; 4c, moderate suspicion; and 5, highly suggestive of malignancy. To allow for meaningful receiver operating characteristic (ROC) analysis, we did not allow use of a 0 BI-RADS score. The ability to recommend additional imaging was separately allowed. Investigators were also asked to rate likelihood of malignancy from 0% to 100% to provide a scale that would potentially improve the ROC analysis. Recommendations for routine annual follow-up, short interval follow-up in 6 months, additional imaging, and biopsy were recorded separately from assessments.

Determination of Reference Standard

Reference standard information is a combination of biopsy results within 365 days and clinical follow-up at 1 year. One year follow-up was targeted for 365 days after the last screening date and very few visits were early; of 2637 participants, 32 (1.2%) occurred before 11 months and 12 (0.46%) before 10.5 months. The absence of a known diagnosis of cancer on a participant interview, review of medical records at the 1-year screening follow-up, or both was considered *disease negative*, as were 3 cases with double prophylactic mastectomies. Biopsy results showing cancer (in situ or infiltrating ductal carcinoma, or infiltrating lobular carcinoma) in the breast or axillary lymph nodes were considered malignant, *disease positive*, as was 1 other invasive cancer, which proved to be a case of melanoma metastatic to axillary lymph nodes. The melanoma case was retained in the analysis because of its classification at the time the database was locked for analysis. Excision was prompted for core biopsy results of atypical or high-risk lesions including atypical ductal or lobular hyperplasia, lobular carcinoma in situ (LCIS), atypical papilloma, and radial sclerosing lesion.

Statistical Considerations

Statistical software used to perform this analysis was SAS, version 9.1 (SAS Institute Inc, Cary, North Carolina), STATA, version 9.2 (STATA Corp, College Station, Texas), S-PLUS, version 7

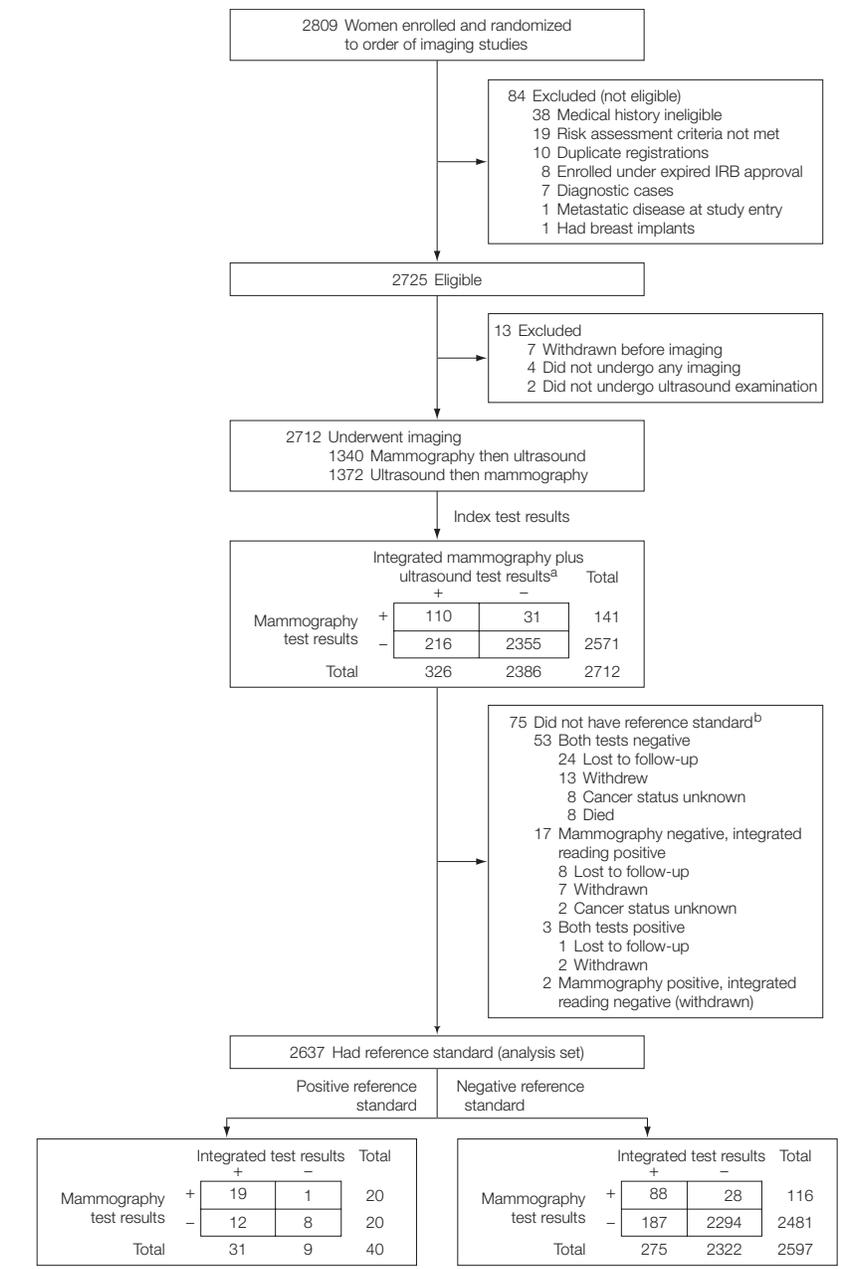
(Insightful Corp, Seattle, Washington), and ROCKIT, version 0.9.4 beta (available from the Kurt Rossmann Laboratories for Radiologic Image Research, University of Chicago, Chicago, Illinois). All *P* values were reported as 2-sided. *P* < .05 was set as the threshold for significance. All confidence intervals (CIs) are reported at the 95% level.

The primary unit of analysis is the participant, with the most severe breast imaging assessment on mammography or on mammography plus ultrasound used as the primary end point. A BI-RADS assessment of 4a, 4b, 4c, or 5 was considered positive (seen and suspicious) for the mammographic or ultrasonographic imaging test or combination of tests, and an assessment of BI-RADS 1, 2, or 3 was considered negative, as is standard in audits of mammographic outcomes.^{6,29} We separately analyzed results based on recommendations, with additional imaging or biopsy or both considered positive and short interval or routine follow-up considered negative. Sample-size projections were designed to achieve both the desired level of statistical precision for estimating the yields and at least 80% power to detect a difference in the yields of at least 3 per 1000, while allowing for 17% missing data.

The diagnostic yield (ie, the proportion of women with a positive screen test and positive reference standard), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were estimated as simple proportions with exact 95% CIs. To account for the natural pairing of assessments within a participant, the McNemar test was used to compare the diagnostic yields, sensitivity, and specificity (TABLE 2) and the test was inverted to provide a CI for their difference. Conditional logistic regression was used where appropriate. Comparison of PPVs and NPVs was done according to Leisenring et al.³⁰ For sensitivity at the lesion level, we accounted for clustering by using a logistic regression with robust SEs. Empirical and model-based ROC curves were estimated from degree of suspi-

cion (BI-RADS) and quasi-continuous probability scales pooled across the study.³¹ The areas under the curve (AUCs) were compared under a bivariate, binormal model that accounts for the paired-test design (ie, every par-

Figure 1. Flowchart of Protocol



IRB indicates institutional review board. *Positive reference standard* is a diagnosis of cancer within 365 days of the initial screening examination. *Negative reference standard* is the absence of a diagnosis of cancer at 1 year follow-up or, for 3 cases, double prophylactic mastectomy. Early second-year screens contribute to the reference standard within the 365-day window.

^aA Breast Imaging Reporting Data System score of 3 or more was considered a positive test result; a score less than 3, negative. One thousand eight hundred thirty participants with both negative mammographic and negative ultrasonographic results were imputed as having a negative integrated reading.

^bBecause of the paired design, missing reference standard data would not bias the comparison of mammography with integrated mammography and ultrasound but may affect generalizability.

participant underwent both screening modalities).^{32,33} The paired-study design eliminates confounding by participant characteristics in the primary comparison between modalities.

Of 2725 eligible participants enrolled, only 3.23% (88) were excluded due to missing data. Thirteen (0.48%) never completed imaging and 75 (2.75%) yielded no reference standard information (Figure 1). The analysis cohort, consisting of all eligible participants with assessment data and reference standard (n=2637), was compared with the full eligible study cohort (n=2725) on baseline characteristics to detect potential biases (Table 1). We note that among the 88 participants with missing data, we would expect only 1 cancer if the data are missing at random.

RESULTS

There were no differences in demographics or risk factors between the analysis cohort of 2637 (4786 breasts) and the overall eligible group of 2725. (Figure 1 and Table 1). The mean (SE) age at enrollment was 55 years (0.2; range, 25-91 years). Fourteen hundred women (53.09%) had a personal history of breast cancer. Nine of 23 women who carried either the *BRCA-1* or *BRCA-2* mutation also had a history of breast cancer, as did 4 of 8 women who had undergone chest or mediastinal radiation. Seventy-three percent of participants had undergone mammography in no less than 11 full months to no more than 14 months before entering the study, 11% had prior screening ultrasound, and 7% had prior contrast-enhanced breast magnetic

resonance imaging at least a year before entering the study.

Forty of 2637 participants (1.5%) were diagnosed with cancer, 39 of whom had breast cancer: 6, DCIS; 20, invasive ductal carcinoma (IDC) with or without DCIS; 3, invasive lobular carcinoma; and 10, mixed invasive ductal and lobular carcinoma with or without DCIS. One participant had melanoma metastatic to axillary nodes with no evidence of cancer in the breasts. One patient with IDC had contralateral DCIS (41 total breasts with cancer). Four patients had multifocal invasive cancer (45 total malignant lesions). Median size of invasive cancers (considering only the largest per participant) was 12.0 mm (range, 4-40 mm; interquartile range [IQR], 8-18 mm; mean [SE], 14 [1.5] mm; 95% CI,

Table 2. Summary of Performance Characteristics of Screening With Combined Mammography Plus Ultrasound Compared With Mammography Alone at the Participant Level^a

	Mammography Plus Ultrasound ^b	Mammography Alone	Comparison of Mammography Plus Ultrasound vs Mammography Alone		Ultrasound Alone ^d
			Difference	P Value ^c	
Yield per 1000					
No./total	31/2637	20/2637			20/2636
% (95% CI)	11.8 (8 to 16.6)	7.6 (4.6 to 11.7)	4.2 (1.1 to 7.2)	.003	7.6 (4.6 to 11.7)
Sensitivity					
No./total	31/40	20/40			20/40
% (95% CI)	77.5 (61.6 to 89.2)	50 (33.8 to 66.2)	27.5 (9.5 to 45.5)	.003	50.0 (33.8 to 66.2)
Specificity					
No./total	2322/2597	2481/2597			2383/2596
% (95% CI)	89.41 (88.16 to 90.57)	95.53 (94.67 to 96.30)	-6.12 (-7.24 to -5)	<.001	91.80 (90.67 to 92.82)
Area under ROC curve					
BI-RADS	0.91 (0.84 to 0.96)	0.78 (0.67 to 0.87)	0.13 (0.04 to 0.22)	.003	0.80 (0.70 to 0.88)
% Probability of malignancy	0.90 (0.83 to 0.95)	0.68 (0.53 to 0.80)	0.23 (0.10 to 0.35)	<.001	0.75 (0.62 to 0.85)
Odds Ratio^e					
Positive predictive value					
No./total	31/306	20/136			20/233
% (95% CI)	10.1 (7.0 to 14.1)	14.7 (9.2 to 21.8)	0.65	.03	8.6 (5.3 to 13.0)
Negative predictive value					
No./total	2322/2331	2481/2501			2383/2403
% (95% CI)	99.61 (99.27 to 99.82)	99.20 (98.77 to 99.51)	2.08	.004	99.17 (98.72 to 99.49)

Abbreviations: BI-RADS, Breast Imaging Reporting Data System; CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristic.

^aEstimates are rounded for presentation. For the comparison column, the counts are replaced by the metric of comparison. Calculations include 1 malignancy that was considered suspicious on mammography and was considered probably benign after integration with ultrasound and 1 malignancy that was not seen on initial imaging that was discovered to be melanoma (from prior primary in back) metastatic to axillary lymph nodes and diagnosed 205 days after study entry.

^bThis table displays screening data. After complete diagnostic workup, results for mammography plus ultrasound are as follows: yield, 31 of 807 participants, 38.4 per 1000 (95% CI, 26.2-54.1); sensitivity, 31 of 35 participants, 88.6% (95% CI, 73.3%-96.8%); specificity, 550 of 772 participants, 71.2% (95% CI, 67.9%-74.4%); PPV, 31 of 253 participants, 12.3%, (95% CI, 8.5%-16.9%); 550 of 554 participants, 99.3% (95% CI, 98.2%-99.8%); area under the curve BI-RADS, 0.91 (95% CI, 0.84-0.95); and area under the curve percentage of probability of malignancy, 0.90 (95% CI, 0.83-0.94).

^cTesting H₀: mammography plus ultrasound is equivalent to mammography alone.

^dUltrasound results were included for completeness. The study was not designed to permit direct comparison of mammography with ultrasound alone. For one participant, the ultrasound alone interpretation was not available.

^eOdds ratios compare the odds of PPV or NPV for mammography plus ultrasound with the odds of PPV or NPV for mammography alone. The method of comparison is given by Leisenring et al,³⁰ which accounts for the paired design but does not readily permit the construction of a CI.

11.1-17.4 mm). Axillary lymph node staging was performed for 25 participants with invasive cancer, with nodal metastases found in 5 (20%, including the melanoma); axillary staging was not performed for 6 participants with recurrent breast cancer nor was it performed for 3 others.

At the participant level, based on BI-RADS assessments, 20 of 40 (50%) of cancers were identified on mammography for a yield of 7.6 per 1000 (Table 2 and TABLE 3); 5 of 6 DCIS lesions (83%) were seen only on mammography. Fifteen invasive cancers, with a median size of 12 mm (range, 4-25 mm; IQR, 7-20 mm; mean [SE], 14 [1.9] mm; 95% CI, 9.9-18.2 mm) were seen on mammography, with axillary nodes negative in 7 of 10 participants (70%) with staging.

Seven invasive cancers were suspicious only on mammography and 8 were suspicious on both mammography and on ultrasound. Ultrasound alone depicted cancer in 12 participants: 1, DCIS, and 11, invasive cancers with median size of 10 mm (range, 5-40 mm; IQR, 6-15 mm; mean [SE] mm, 12.6 [3.0]; 95% CI, 6.0-19.1 mm), with axillary nodes negative in 8 of 9 participants (89%) with staging. One 4-mm IDC lesion considered suspicious initially on mammography (true positive on mammography) was downgraded to a BI-RADS score of 3 after integration with ultrasound (false negative on mammography plus ultrasound), even though it was still recalled for additional mammographic views (thought to be probably benign after recall, and benign at the

6-month follow-up), and was diagnosed when the patient presented with palpable metastatic adenopathy 264 days after study entry. This is not included among interval cancers.

Thirty-one cancers were depicted in 2637 participants by the combination of mammography plus ultrasound, producing a yield of 11.8 per 1000 women and an increased yield due to ultrasound of 4.2 per 1000 (95% CI, 1.1-7.2; Table 2) over mammography alone. The diagnostic accuracy of mammography alone was 0.78 (95% CI, 0.67-0.87), for ultrasound alone was 0.80 (95% CI, 0.70-0.88), and for combined mammography plus ultrasound was 0.91 (95% CI, 0.84-0.96, Table 2, FIGURE 2). The AUC for mammography plus ultrasound did not change

Table 3. Sensitivity, Specificity, and Positive Predictive Value by Participant for Mammography Alone vs Combined Mammography Plus Ultrasound

BI-RADS Score	Mammography Alone								Mammography Plus Ultrasound							
	Total No.		Cumulative No.		Sensitivity of Detecting Cancer, %		Specificity, %	PPV Level, %	Total No.		Core or Surgical Biopsy		Cyst Aspiration Only		Atypical Biopsy	
	Breast Women	Cancer	Women	True Positives	All	Invasive			Women ^a	Lesions	Women	Lesions	Women ^b	Lesions	Women	Lesions
5, Highly suggestive of malignancy	5	4	5	4	10.0	8.8	99.96	80.0	5	5	0	0	0	0		
4C, Moderate suspicion	9	4	14	8	20.0	20.6	99.77	57.1	9	12	0	1	1	1		
4B, Intermediate suspicion	32	7	46	15	37.5	32.4	98.81	32.6	22	24	1	2	0	0		
4A, Low suspicion	90	5 ^c	136	20 ^c	50.0	44.1	95.53	14.7	33	43	3	6	2	2		
3, Probably benign	177	1	313	21	52.5	47.1	88.76	6.7	16	19	12	16	0	0		
2, Benign ^d	1421	14	1734	35	87.5	88.2	34.58	2.0	112	131	37	59	7	7		
1, Negative	903	5	2637	40	100	100	0	1.5	70	85	18	29	6	8		
Mammography Plus Ultrasound																
5, Highly suggestive of malignancy	6	5	6	5	12.5	11.8	99.96	83.3	6	8	0	1	0	0		
4C, Moderate suspicion	26	12	32	17	42.5	47.1	99.42	53.1	26	35	0	0	2	2		
4B, Intermediate suspicion	68	10	100	27	67.5	67.6	97.19	27.0	58	73	5	9	4	4		
4A, Low suspicion	206	4	306	31	77.5	73.5	89.41	10.1	112	128	42	74	6	6		
3, Probably benign	401	4 ^c	707	35 ^c	87.5	85.3	74.12	5.0	40	49	17	22	3	5		
2, Benign	98	0	805	35	87.5	85.3	70.35	4.3	3	3	1	1	0	0		
1, Negative ^d	1832 ^e	5	2637	40	100	100	0	1.5	22	23	6	6	1	1		

Abbreviations: BI-RADS, Breast Imaging Reporting Data System.

^aIncludes 1 prophylactic mastectomy with ductal carcinoma in situ in a patient with contralateral invasive ductal cancer, but does not include 3 negative double-prophylactic mastectomies nor 5 other contralateral prophylactic mastectomies.

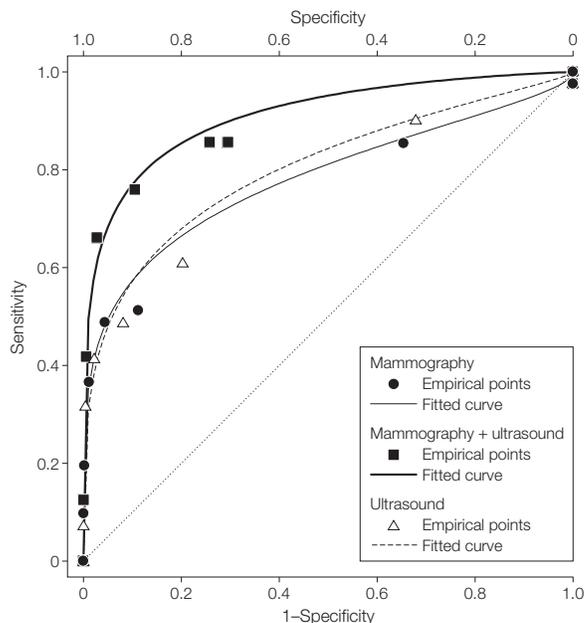
^bAtypical includes atypical ductal hyperplasia or lobular hyperplasia, lobular carcinoma in situ, atypical papilloma, and radial sclerosing lesion

^cOne 4-mm invasive ductal cancer was considered BI-RADS 4a on mammography, missed on ultrasound, and classified as BI-RADS 3 on integrated mammography plus ultrasound. This was diagnosed when the participant presented with palpable metastatic adenopathy 264 days after study entry.

^dThe participant with cancer due to melanoma presenting with axillary metastases was classified as BI-RADS 2 on mammography and BI-RADS 1 on integrated mammography plus ultrasound.

^eInvestigators were not asked to integrate findings on mammography plus ultrasound if both examination results recommended routine follow-up. These cases were analyzed as having negative results on mammography plus ultrasound. Of 1832 cases analyzed as BI-RADS 1 on mammography plus ultrasound, 342 were BI-RADS 2 on mammography, 318 were BI-RADS 2 on ultrasound, 747 were BI-RADS 2 on both; 2 were BI-RADS 3 on initial mammography and downgraded to BI-RADS 1 after integration with ultrasound; the remaining 423 were BI-RADS 1 on both mammography and ultrasound.

Figure 2. Sensitivity and Specificity of Mammography Plus Ultrasound in Detecting Breast Cancer



Receiver operating characteristic (ROC) curves were calculated based on a bivariate, binomial model (See "Methods" section for details). Table 2 presents summary characteristics for these curves. The ultrasound ROC is included for completeness; the study was not designed to permit direct comparison to ultrasound alone. The fitted area under the curve for mammography alone is 0.78 (95% confidence interval [CI], 0.67-0.87); for mammography plus ultrasound, 0.91 (95% CI, 0.84-0.96); and for ultrasound alone, 0.80 (95% CI, 0.70-0.88).

when incorporating full diagnostic workup that included additional mammographic views.

Defined as the percentage of participants with a BI-RADS 4a assessment or higher, without a diagnosis of cancer in the following 12 months, the false-positive rate for mammography alone was 4.4% (136 had a BI-RADS score of 4a or higher, of whom 20 had cancer: 116 were false positive of 2637 participants [95% CI, 3.7%-5.3%; Table 3]); for ultrasound alone, the false-positive rate was 8.1% (213 of 2637; 95% CI, 7.1%-9.2%); and for combined mammography plus ultrasound was 10.4% (275 of 2637; 95% CI, 9.3%-11.7%). In 5.2% of participants (136 of 2637; 95% CI, 4.3%-6.1%), ultrasound, but not mammography, resulted in a suspicious assessment and biopsy and 8.8% (12 of 136; 95% CI, 4.6%-14.9%) of these participants had cancer. Seventy-one participants had only a cyst aspiration, without a bi-

opsy, with no malignancies among these lesions; 43 of these participants had a suspicious assessment only on ultrasound and 2 had a suspicious assessment only on mammography.

TABLE 4 details the recommendations by modality. To calculate the PPV1⁶ of recall, the number of participants with cancer are divided by those who were recalled for additional evaluation, biopsy, or both. For mammography, 21 participants were diagnosed with cancer among of 276 participants who underwent additional evaluation or biopsy, or both for a PPV1 of 7.6% (95% CI, 4.8%-11.4%); for ultrasound the PPV1 was 22 of 337 (6.5%; 95% CI, 4.1%-9.7%); and for combined integrated mammography plus ultrasound, 32 of 436, 7.3% (95% CI, 5.1%-10.2%). Of those 276 participants recalled from routine mammography, after complete diagnostic workup, 84 participants were recommended for biopsy, of whom 19 had cancer, resulting in a PPV2⁶ of 22.6% (95% CI, 14.2%-33%). For ultra-

sound, 21 of 235 patients with a biopsy recommendation after workup had cancer, resulting in a PPV2 of 8.9% (95% CI, 5.6%-13.3%). Although 1 of these cancers was classified as BI-RADS 3 based on the initial ultrasound, it was nevertheless worked up and classified as BI-RADS 4b on mammography. After mammography plus ultrasound and a full diagnostic workup, 31 of 276 participants who had undergone biopsy had cancer, resulting in a PPV2 of mammography plus ultrasound of 11.2% (95% CI, 7.8%-15.6%).

Short Interval Follow-up and BI-RADS 3

Based on mammographic results, 177 women (6.7%) were classified as BI-RADS 3 (Table 3); of those, 1 (0.6%) was diagnosed with cancer detected at the early second screen, 363 days after study entry (after initial additional mammographic recall at time 0 for unrelated benign findings). Three hundred twenty-one participants (12.2%) were classified as BI-RADS 3 based on ultrasonographic screening, 5 of whom (1.6%) were diagnosed with cancer within the first 12 months of follow-up. Of the 5 participants with cancer who were classified as BI-RADS 3 on ultrasound, 3 were BI-RADS 5 on mammography and were diagnosed from 1 to 23 days after initial screens were completed. Two women had interval cancers that were identified incidentally as a result of a 6-month follow-up ultrasound for complicated cysts (the first was a 7-mm IDC found at surgery in adjacent tissue after a core biopsy result of LCIS from the lesion being followed up; the other was a 27-mm IDC-DCIS adjacent to the cyst being followed up). Both participants were node negative.

Based on results from mammography, short-interval follow-up was recommended for 59 (2.2%) of 2637 participants (95% CI, 1.7%-2.9%), and based on ultrasound, recommendations for short-term follow-up were made for 227 (8.6%) participants (95% CI, 7.6%-9.7%). Two hundred twenty of these recommendations were based on ultrasound alone. Two hundred eighty-six participants (10.8%) were

recommended for short-term follow-up after mammography plus ultrasound (95% CI, 9.7%-12.1%).

Initial assessments of 27 participants as BI-RADS 3, 7 as BI-RADS 4a, and 1 as BI-RADS 4b based on mammography were downgraded to BI-RADS 2 after integrating mammographic and ultrasonographic results. Similarly, initial assessments of 26 participants as BI-RADS 3, 3 as BI-RADS 4a, 4 as BI-RADS 4b, and 1 as BI-RADS 5 based on ultrasound were downgraded to a BI-RADS score of 2 after integrating ultrasonographic and mammographic results.

Interval Cancers

Eight participants had cancer not considered suspicious on either mammography or ultrasound, with cancer identified during the 12 months after initial screening, ie, *interval cancers*. Three node-negative cancers (an 8-mm IDC and ILC, a 35-mm ILC, and a 20-mm IDC-DCIS) were identified at the second screen (performed early, after 11 full months), with biopsies taken from 359 to 364 days after study entry. One participant noted a palpable lump, with biopsy showing a 12-mm mixed IDC/ILC 337 days after study entry. One participant presented with skin recurrence of prior breast cancer 231 days after study entry. Two cancers were found at the 6-month follow-up ultrasound as detailed in the section on short-interval follow-up. One non-breast malignancy was identified in the interval in a participant with prior melanoma of the back, who, 6 years later, developed a palpable axillary mass due to metastatic adenopathy, with no evidence of malignancy within the breasts. Thus, the interval cancer rate was 8 of 40 (20%) if the melanoma case is included as cancer, or 7 of 39 (18%) if not; only 2 of 39 participants (5.1%) with breast cancer were identified because of symptoms in the interval between screenings—or 3 of 39 (7.8%), if one includes the 4-mm IDC seen on initial mammography but not on additional imaging or at the 6-month follow-up, but which was diagnosed when the par-

Table 4. Summary of Recommendations After Screening Mammography, Ultrasound, Combined Mammography Plus Ultrasound, and After Diagnostic Workup

Recommendation	No. of Cancers/No. of Participants With This Recommendation (%)			
	Mammography	Ultrasound	Mammography Plus Ultrasound	After Additional Evaluation
Additional imaging	15/249 (6.0) ^a	4/98 (4.1) ^b	8/175 (4.6)	0/6
Biopsy ^c	4/21 (19.0)	12/195 (6.2)	20/239 (8.4)	31/272 (11.4)
Additional imaging and biopsy	2/6 (33.3)	6/44 (13.6)	4/22 (18.2)	0/4
Short-interval follow-up	0/59	3/227 (1.3)	3/286 (1.0)	3/336 (0.9)
Annual follow-up ^d	19/2302 (0.8)	15/2072 (0.7)	5/1915 (0.3)	6/2019 (0.3)
Missing		0/1		
Total	40/2637 (1.5)	40/2637 (1.5)	40/2637 (1.5)	40/2637 (1.5)

^aOne participant with a 4-mm invasive ductal cancer was given a 4a Breast Imaging Reporting Data System (BI-RADS) score on mammography and was recommended for additional imaging. Ultrasound was negative, and the participant was given a BI-RADS score of 3 with recommendation for short-interval follow-up on integrated mammography plus ultrasound. Cancer was diagnosed when the participant presented with palpable metastatic adenopathy 264 days after study entry. Another participant was recalled for additional imaging of calcifications thought to be a BI-RADS score of 3 on screening and benign on additional imaging, with cancer diagnosed 363 days after study entry due to a finding on the (early) 12-month screen.

^bTwo participants with cancer were classified as BI-RADS 3 on ultrasound with a recommendation for additional imaging. One had a BI-RADS 4b lesion on mammography and was biopsied. The other was recommended for short-interval follow-up after integrated mammography plus ultrasound and was diagnosed at the 6-month follow-up ultrasound.

^cBiopsy includes cyst aspirations for diagnostic uncertainty: 1 prompted by mammography and ultrasound; 46 only by ultrasound; 52 after combined mammography plus ultrasound integrated interpretation, and 49 after any additional evaluation was completed.

^dThe participant with cancer due to axillary metastasis from melanoma was recommended for routine follow-up on all imaging.

participant presented with palpable metastatic adenopathy 264 days after study entry. A ninth breast had cancer not seen on either mammography or ultrasound: DCIS was identified only at prophylactic mastectomy after diagnosis of contralateral multifocal IDC seen only on ultrasound.

Cancers seen only on ultrasound were evenly distributed across breast density categories (TABLE 5). The data were inconclusive with respect to most differences between film-screen and digital mammography; however, slightly higher specificity was observed with digital mammography than with film screen (97.0% vs 94.7%, $P = .007$).

In 1400 women with a personal history of breast cancer, 28 (2.0%) were found to have cancer, with 9 of 28 (32%) seen only on ultrasound (Table 5). Cancers were evenly distributed between the breast ipsilateral to the initial cancer and contralateral disease. Among 1237 women with risk factors other than a personal history of breast cancer, 12 (1.0%) were found to have cancer, 3 of which cancers (25%) were seen only on ultrasound. Significantly more cancers overall were found

in women with a personal history of cancer ($P = .03$), but there was no difference in supplemental yield of ultrasound in women with or without a personal history of breast cancer.

The median time to perform screening breast ultrasound was 19 minutes (range, 2-90; IQR, 12-27, mean [SE], 20.8 [0.3], 95% CI, 20.3-21.3 minutes) for a bilateral scan and 9 minutes for a unilateral scan (range, 1 to 70; IQR, 5-15; mean [SE], 11.6 [0.4], 95% CI, 10.7-12.4 minutes). A median of another 2.0 minutes was spent in the room with the participant (range, 0-19; IQR, 2-3; mean [SE], 2.7 [0.04]; 95% CI, 2.6-2.7 minutes). For 869 (33.0%) of 2637 participants, the investigator scanned at least 1 axilla while performing ultrasonographic scanning of the breast(s). Ninety-four percent of breasts were less than 4 cm thick.

COMMENT

Supplemental physician-performed screening ultrasound increases the cancer detection yield by 4.2 cancers per 1000 women at elevated risk of breast cancer, as defined in this protocol (95% CI, 1.1-7.2 cancers per 1000) on a single, prevalent screen. This is similar to rates

of ultrasound-only cancers of 2.7 to 4.6 cancers per 1000 women screened in other series.^{8,13-17,24} As in prior studies, the vast majority of cancers seen only on ultrasound were invasive because DCIS is difficult to see on ultrasound. All but 1 cancer seen only on ultrasound was node negative. Invasive cancers not seen on mammography can be expected to present as interval cancers with a worse prognosis: detection of asymptomatic, mammographically occult, node-negative invasive carcinomas with ultrasound should reduce mortality from breast cancer, although mortality was not an end point of this study.

Strengths of our study include its matching within a participant, and examinations performed by radiologists who were masked to results

of the other examination. Randomized order of these tests helped control biases of recruiting women with vague mammographic abnormalities. Furthermore, these results were consistent and generalizable across 21 international centers. The radiologist investigators in this trial were all specialists in breast imaging who met experience requirements and completed qualification tasks. As such, our results may vary slightly from those observed in general practice, even though similar results were observed by Kaplan¹⁶ for which study technologists performed screening ultrasound. Educational materials used for radiologist investigator training in ultrasound lesion detection and characterization are archived by ACRIN.

The use of the Gail and Claus models to calculate risk may have affected the racial distribution of participants, for the Gail model is known to underestimate risk in African Americans.³⁴ Neither model has been validated in other races other than whites,^{34,35} although Gail et al³⁶ have recently validated a new risk assessment tool based on data from the Contraceptives and Reproductive Experiences (CARE) Study, which involved African American women (which was not available for use in this protocol).

In our elevated-risk study population, enriched in women with dense breasts, mammographic sensitivity was only 50% (95% CI, 33.8%-66.2%) and the sensitivity of mammography plus ultrasound was 77.5% (95% CI, 61.6%-

Table 5. Summary of Cancers Identified by Participant as a Function of Breast Density, Mammogram Type, and Eligibility Risk Factor

	No. (%) of Participants in the Analysis Set (n = 2637)	No. of Cancers (No. Invasive)				
		All Cancers	BI-RADS Score >3			BI-RADS Score ≤3, Mammography Plus Ultrasound ^b
			Mammography Alone ^a	Ultrasound Alone	Mammography Plus Ultrasound	
Total	2637 (100.0)	40 (34)	12 (7)	12 (11)	8 (8)	8 (8)
Breast density, % ^c						
≤25	53 (2.0)					
26-40	297 (11.3)	6 (4)	3 (2)	3 (2)		
41-60	811 (30.7)	14 (13)	4 (3)	4 (4)	3 (3)	3 (3)
61-80	968 (36.7)	11 (9)	3 (1) ^a	2 (2)	4 (4)	2 (2)
>80	506 (19.2)	9 (8)	2 (1)	3 (3)	1 (1)	3 (3) ^b
Missing	2 (0.1)					
Total	2637 (100.0)	40 (34)	12 (7)	12 (11)	8 (8)	8 (8)
Mammogram type						
Digital	923 (35.0)	14 (14)	3 (3)	4 (4)	3 (3)	4 (4)
Film-screen	1714 (64.9)	26 (20)	9 (4) ^a	8 (7)	5 (5)	4 (4) ^b
Total	2637 (100.0)	40 (34)	12 (7)	12 (11)	8 (8)	8 (8)
Eligibility risk factor						
Personal history of breast cancer	1400 (53.1)	28 (25)	7 (5)	9 (8)	7 (7)	5 (5) ^b
Lifetime risk ≥25%	497 (18.8)	5 (4)	3 (2) ^a	1 (1)		1 (1)
5-y risk, Gail model ≥2.5%	403 (15.3)	4 (4)		2 (2)	1 (1)	1 (1)
≥1.7% and extremely dense breasts	223 (8.5)	2 (1)	1 (0)			1 (1)
ADH/ALH/LCIS or atypical papilloma	83 (3.1)	1 (0)	1 (0)			
Mutation in <i>BRCA1</i> or <i>BRCA2</i> genes	23 (0.9)					
History of prior chest and/or mediastinal and/or axillary irradiation	8 (0.3)					
Total	2637 (100.0)	40 (34)	12 (7)	12 (11)	8 (8)	8 (8)

Abbreviations: ADH, atypical ductal hyperplasia; ALH, atypical lobular hyperplasia; LCIS, lobular carcinoma in situ.

^aOne invasive ductal cancer seen on initial mammography was classified as probably benign (BI-RADS 3) after integrated interpretation with ultrasound and had breast density between 41% and 60%, mammographic type film-screen, and lifetime risk of 25% or higher.

^bMelanoma case had breast density of more than 80%, mammographic type film-screen, and personal history of breast cancer.

^cBreast density is overall visually estimated breast density. All participants were visually judged to have at least heterogeneously dense parenchyma in at least 1 quadrant of the breast(s) on prior mammography, except for 43 participants who had no prior mammography.

89.2%; Table 2). From a detection standpoint, it may be reasonable to offer supplemental screening ultrasound to women with similar risk criteria. As stated, dense breast tissue is common: approximately half of women younger than 50 years and a third of older women have dense breast parenchyma.⁵ Approximately 6% of women presenting for routine annual mammography have a personal history of breast cancer,²⁹ and 15% have a family history of breast cancer.²⁹

Our ongoing study, allowing for contrast-enhanced breast magnetic resonance imaging (MRI) within 8 weeks of the final 24-month mammography and ultrasound screening round, may soon shed some light on the possible competitive roles of ultrasound and MRI as adjuncts to mammographic screening for breast cancer. Across 4 other series for which screening mammography, ultrasound, and MRI had been performed for women at very high risk of breast cancer, the combined sensitivity of mammography and ultrasound averaged 55% vs 93% after combined mammography and MRI.³⁷⁻⁴⁰ There appears to be no role for screening ultrasound in women undergoing screening MRI, even though ultrasound may be helpful in guiding biopsy of suspicious findings seen first on MRI.³⁷⁻⁴⁰ Ultrasound may be more appropriate than MRI for screening women of intermediate risk due to its reduced cost relative to MRI. Many of the cancers seen only on MRI are small, node-negative invasive cancers.³⁷⁻⁴⁰ Unlike ultrasound, MRI readily depicts DCIS,⁴¹ although DCIS remains overrepresented among false-negative MRI examinations.⁴² It is uncertain whether detection of DCIS is required or whether detection of node-negative invasive breast cancer is sufficient for a screening test. It will be important to see the stage distribution of breast cancers in subsequent rounds of screening with mammography plus ultrasound in this study and to know how many invasive cancers will be seen only on MRI at the 24-month time point.

Despite a 20% interval cancer rate (8 of 40 participants with cancer) in our

series, none of the interval breast carcinomas were node positive; the only interval cancer that was node positive was a nonbreast cancer (melanoma metastatic to axillary nodes). Another cancer considered suspicious on initial mammography (and therefore not included among "interval cancers") was considered probably benign after full diagnostic workup and went unbiopsied until the patient presented with palpable, metastatic nodes, yet was only 4 mm in size at eventual detection. One interval cancer was a skin recurrence of prior breast cancer.

Ultrasound is well tolerated, the technology is widely available, and it does not require intravenous contrast material. If, however, screening ultrasound is to be widely implemented, several major issues remain. First, it will be very important to know the role of annual screening ultrasound in addition to mammography, and such a study is in progress with participants in this protocol. The time to perform bilateral screening ultrasound is problematic, at a median of 19 minutes. This does not include comparison to prior studies, discussion of results with patients, nor creation of a final report, although the time may be artificially prolonged by protocol requirements to measure each lesion other than a simple cyst in 2 planes and to fully characterize each such lesion with and without spatial compounding and with and without color or power Doppler. Nineteen minutes is considerably longer than the average 4 minutes 39 seconds reported by Kolb et al⁸ for physicians scanning or the average 10 minutes reported by Kaplan¹⁶ for technologists. Currently, there is only a single billing code for breast ultrasound (current procedural terminology code 76645), and Medicare global reimbursement is \$85 in 2008, which does not fully cover the costs of performing and interpreting the examination. Outcomes similar to those of our physician-performed study have been reported with technologist-performed ultrasound,¹⁶ and specialized training of technologists is encouraged to

counter a current shortage of qualified physician and technologist personnel. Further validation of technologist-performed screening breast ultrasound is encouraged. Automated whole-breast ultrasound may facilitate implementation and profitability of screening ultrasound but will result in hundreds of images to be reviewed and stored, with attendant increased capital and professional costs and potential increased malpractice exposure; validation of such methods is needed. The full costs of screening breast ultrasound in this protocol, including the costs of induced additional testing and biopsy, are being analyzed and reported separately.

The final barrier to implementing screening ultrasound is the risk of false-positive results. The performance characteristics of mammography were within accepted ranges (10.5% recalled for additional imaging or biopsy; 3.2% of participants biopsied after full workup, with 23% proving malignant; 2.2% recommended for short interval follow-up). We observed a 5.4% recall rate for ultrasound (142 of 2637 recommended for additional imaging), which may be artificially low in this series because physicians performed the screening ultrasound and could directly evaluate lesions in real-time. Of 2637 participants, 233 (8.8%) participants had findings considered suspicious on ultrasound with 136 participants having suspicious findings on ultrasound but not mammography and prompting biopsy, and 235 participants (8.9%) were recommended for biopsy based on ultrasound after full workup. Only 20 of 233 (8.6%) of participants with suspicious ultrasonographic findings—12 (8.8%) of 136 of those with suspicious findings biopsied based on ultrasound alone—and 21 of 235 (8.9%) of participants whose lesions were recommended for biopsy based on ultrasound proved to have cancer. The 8.8% to 8.9% PPV of biopsies prompted by ultrasound in our study is similar to the 11% rate seen across prior series.^{20,43} Diagnostic uncertainty for complicated

cysts remains a major source of false-positive results, with 43 participants undergoing only cyst aspiration included among those with a suspicious finding on ultrasound. Elastography, in which the deformability of the mass is assessed during ultrasound, can help distinguish complicated cysts from suspicious solid masses and should reduce this source of false positives.⁴⁴ Another 227 participants (8.6%) were recommended for short interval follow-up based on ultrasound, similar to the 6.3% rate across other series.^{8,15,16,45} Whether the risk of false-positive results with ultrasound will diminish in our study population with subsequent screening rounds, as has been seen with mammography⁴⁶ and in small series with both ultrasound and MRI³⁷ is under evaluation. We have been separately quantifying patient anxiety and discomfort (ie, "process utility"⁴⁷) induced by addition of screening ultrasound.

CONCLUSION

The addition of a single screening ultrasonographic examination to mammography for women at elevated risk of breast cancer results in increased detection of breast cancers that are predominantly small and node-negative. We defined *elevated risk* using a variety of criteria, including personal history of breast cancer, prior atypical biopsy, and elevated risk by Gail or Claus models or both. Recent literature⁴³ suggests that any combination of factors that confers 3-fold relative risk compared with women without the risk factor would be "high risk," including dense breast tissue.⁹ Across all series to date, over 90% of cancers seen only on ultrasound have been in women with more than 50% dense breast tissue,^{20,24} although 3 of 12 cancers (25%) seen only on ultrasound in this series were in women with only 26% to 40% dense breast tissue (as visually estimated), suggesting that women with other risk factors may benefit from screening ultrasound even if their breast tissue is less dense. The age at which to begin screening women at increased risk would reasonably derive from the age at

which the risk of breast cancer is equal to that for an average woman aged 40 or 50 years, depending on national policy.⁹

The detection benefit of a single screening ultrasound in women at elevated risk of breast cancer is now well validated. However, it comes with a substantial risk of false-positive results (ie, biopsy with benign results and/or short interval follow-up). Our results should be interpreted in the context of recent guidelines recommending annual MRI in women at very high risk of breast cancer.²⁵ Importantly, evaluation of annual (incidence) screening ultrasound is continuing in ACRIN 6666, as is evaluation of a single screening MRI in these women.

Author Affiliations: American Radiology Services Inc, Johns Hopkins Green Spring, Lutherville, Maryland (Dr Berg); Center for Statistical Sciences, Brown University, Providence, Rhode Island (Drs Blume and Cormack and Ms Marques); Feinberg School of Medicine, Northwestern University, Chicago, Illinois (Dr Mendelson); CERIM, Buenos Aires, Argentina (Dr Lehrer); Weinstein Imaging Associates, Pittsburgh, Pennsylvania (Dr Böhm-Vélez); University of North Carolina School of Medicine, Chapel Hill (Dr Pisano); University of Toronto, Sunnybrook and Women's Hospital, Toronto, Ontario, Canada (Dr Jong); University of Texas Southwestern Medical Center, Dallas (Dr Evans); Mayo Clinic, Rochester, Minnesota (Dr Morton); University of Cincinnati, Cincinnati, Ohio (Dr Mahoney); Keck School of Medicine, University of Southern California, Los Angeles (Dr Larsen); Forum Health, Western Reserve Care System, Youngstown, Ohio (Dr Barr); Mallinckrodt Institute of Radiology, Washington University School of Medicine, St Louis, Missouri (Dr Farria); American College of Radiology, Philadelphia, Pennsylvania (Ms Boparai).

Author Contributions: Dr Blume had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Berg, Blume, Mendelson, Pisano.

Acquisition of data: Berg, Blume, Mendelson, Lehrer, Böhm-Vélez, Pisano, Jong, Evans, Morton, Mahoney, Hovanessian-Larsen, Barr, Farria, Boparai.

Analysis and interpretation of data: Berg, Blume, Cormack, Mendelson, Pisano, Marques.

Drafting of the manuscript: Berg, Blume, Cormack, Marques.

Critical revision of the manuscript for important intellectual content: Berg, Blume, Mendelson, Lehrer, Böhm-Vélez, Pisano, Jong, Evans, Morton, Mahoney, Hovanessian-Larsen, Barr, Farria, Boparai.

Statistical analysis: Blume, Cormack, Marques.

Obtained funding: Berg, Blume.

Administrative, technical, or material support: Berg, Mendelson, Lehrer, Pisano, Jong, Morton, Barr, Boparai.

Study supervision: Berg, Blume.

Financial Disclosures: Dr Berg reports that she has served as a consultant to Naviscan PET Systems, MediPattern, and Siemens and has received equipment support from Siemens and a travel grant from General Electric. Dr Mendelson reports that she is a member of the scientific advisory boards of MediPattern and Siemens and has received equipment support from Philips. Dr Böhm-

Vélez reports that she is a member of the physicians advisory board of MediPattern. Dr Pisano reports that her laboratory receives research support from General Electric, Hologic, Konica, Sectra, and Siemens. Dr Jong reports that she has a research collaboration with General Electric. Dr Evans reports that he is a member of the scientific advisory board of Hologic. Dr Mahoney reports that she is a consultant to Johnson & Johnson and SenoRx. Dr Larsen reports that she receives equipment support from Naviscan PET Systems. Dr Barr reports that he is a member of the ultrasound advisory boards of and has received equipment support from Siemens and Philips. The remaining coauthors report no financial disclosures.

ACRIN 6666 Site Investigators: Allegheny-Singer Research Institute, Pittsburgh, Pennsylvania: William R. Poller, MD, principal investigator (PI), Michelle Huerbin, research associate (RA); American Radiology Services—Johns Hopkins Green Spring, Baltimore, Maryland: Wendie A. Berg, MD, PhD (PI), Barbara E. Levit, RT (RA); Beth Israel Deaconess Medical Center, Boston, Massachusetts: Janet K. Baum, MD, and Valerie J. Fein-Zachary, MD (PIs), Suzette M. Kelleher, BA (RA); CERIM, Buenos Aires: Daniel E. Lehrer, MD (PI), Maria S. Ostertag (RA); Duke University Medical Center, Durham, North Carolina: Mary Scott Soo, MD (PI), Brenda N. Prince, RT (RA); Mayo Clinic, Rochester, Minnesota: Marilyn J. Morton, DO (PI), Lori M. Johnson, AAS (RA); Feinberg School of Medicine, Northwestern University, Chicago, Illinois: Ellen B. Mendelson, MD (PI), Marysia Kalata, AA (RA); Radiology Associates of Atlanta, Atlanta, Georgia: Handel Reynolds, MD (PI), Y. Suzette Wheeler, RN, MSHA (RA); Radiology Consultants/Forum Health, Youngstown, Ohio: Richard G. Barr, MD, PhD (PI), Marilyn J. Mangino, RN (RA); Radiology Imaging Associates, Denver, Colorado: A. Thomas Stavros, MD (PI), Margo Valdez (RA); Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada: Roberta A. Jong, MD (PI), Julie H. Lee, BSC (RA); Thomas Jefferson University Hospital, Philadelphia, Pennsylvania: Catherine W. Piccoli, MD, and Christopher R. B. Merritt, MS, MD (PIs), Colleen Dascenzo (RA); David Geffen School of Medicine at University of California Los Angeles Medical Center, Los Angeles: Anne C. Hoyt, MD (PI), Roslynn Marzan, BS (RA); University of Cincinnati Medical Center, Cincinnati, Ohio: Mary C. Mahoney, MD (PI), Monene M. Kamm, AS (RA); University of North Carolina, Chapel Hill: Etta D. Pisano, MD (PI), Laura A. Tuttle, MA (RA), Keck School of Medicine, University of Southern California, Los Angeles: Linda Hovanessian Larsen, MD (PI), Christina E. Kiss, AA (RA); University of Texas M. D. Anderson Cancer Center, Houston: Gary J. Whitman, MD (PI), Sharon R. Rice, AA (RA); University of Texas Southwestern Medical Center, Dallas: W. Phil Evans, MD (PI), Kimberly T. Taylor, AA (RA); Washington University School of Medicine, St. Louis, Missouri: Dione M. Farria, MD, MPH (PI), Darlene J. Bird, RT, AS (RA); and Weinstein Imaging Associates, Pittsburgh, Pennsylvania: Marcela Böhm-Vélez, MD, (PI), Antoinette Cockroft (RA).

Funding/Support: The study was funded by the Avon Foundation and grants CA 80098 and CA 79778 from the National Cancer Institute.

Role of the Sponsors: The Avon Foundation was not involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. The trial was conducted by the American College of Radiology Imaging Network, a member of the National Cancer Institute's Clinical Trials Cooperative Groups Program, and was developed and carried out adhering to the standard cooperative group processes. These processes include review of and input about the trial design from the NCI's Cancer Therapy Evaluation Program (CTEP). Upon CTEP's approval of the research protocol, the NCI was not in-

involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

Additional Contributions: We thank Amanda M. Adams, MPH, Center for Statistical Sciences, Brown University, Providence, Rhode Island, for assistance with data analysis; Eric A. Berns, PhD, University of Colorado, Denver, for ultrasound quality assurance; Cynthia B. Olson, MBA, MHS, and Sophia Sabina, MBA, American College of Radiology (ACR), Philadelphia, Pennsylvania, for administrative assistance; Glenna J. Gabrielli, BS, Stephanie Clabo, BS, CCRP, Jillene DeBari, BA, and Judy M. Green, RT(M) at ACR for data management; Cheryl L. Crozier, RN, ASQ, CQA, and Josephine Schloesser, AS, RT(R)(M), CCRP at ACR for monitoring; Anthony M. Levering, AS, RT(R)(CT) (MR) at ACR for image management; and Nancy S. Fredericks, MBA, at ACR for communications support. We also thank Cecilia M. Brennecke, MD, and other colleagues at American Radiology Services, Johns Hopkins Green Spring, for their support; and Mark D. Schleinitz, MD, MS, at Brown University, Barbara K. LeStage, BS, MHP, consultant to the ACR, Edward A. Sickles, MD, University of California, San Francisco Medical Center, and Elizabeth A. Patterson, MD, Seattle, Washington, for engaging in helpful discussions. We are indebted to the many investigators, coinvestigators, and research associates at the clinical sites. No one was compensated beyond their usual salary for their efforts for this study.

REFERENCES

- Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the US Preventive Services Task Force. *Ann Intern Med*. 2002;137(5 part 1):347-360.
- Tabár 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*. 2000;38(4):625-651.
- Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: what have we learned? *Radiol Clin North Am*. 2004;42(5):793-806.
- Tabár L, Tot T, Dean PB. Early detection of breast cancer: large-section and subgross thick-section histologic correlation with mammographic appearances. *Radiographics*. 2007;27:S5-S35.
- Stomper PC, D'Souza DJ, DiNitto PA, Arredondo MA. Analysis of parenchymal density on mammograms in 1353 women 25-79 years old. *AJR Am J Roentgenol*. 1996;167(5):1261-1265.
- D'Orsi CJ, Bassett LW, Berg WA, et al. *Breast Imaging Reporting and Data System, BI-RADS: Mammography*. 4th ed. Reston, VA: American College of Radiology; 2003.
- Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. *J Natl Cancer Inst*. 2000;92(13):1081-1087.
- 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*. 2002;225(1):165-175.
- Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med*. 2007;356(3):227-236.
- Harvey JA, Bovbjerg VE. Quantitative assessment of mammographic breast density: relationship with breast cancer risk. *Radiology*. 2004;230(1):29-41.
- Pisano ED, Gatsonis C, Hendrick E, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med*. 2005;353(17):1773-1783.
- Buist DS, Porter PL, Lehman C, Taplin SH, White E. Factors contributing to mammography failure in women aged 40-49 years. *J Natl Cancer Inst*. 2004;96(19):1432-1440.
- Buchberger W, Niehoff A, Obrist P, DeKoekoek-Doll P, Dunser M. Clinically and mammographically occult breast lesions: detection and classification with high-resolution sonography. *Semin Ultrasound CTMR*. 2000;21(4):325-336.
- Crystal P, Strano SD, Shcharynski S, Koretz MJ. Using sonography to screen women with mammographically dense breasts. *AJR Am J Roentgenol*. 2003;181(1):177-182.
- Gordon PB, Goldenberg SL. Malignant breast masses detected only by ultrasound: a retrospective review [see comments]. *Cancer*. 1995;76(4):626-630.
- Kaplan SS. Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue. *Radiology*. 2001;221(3):641-649.
- Leconte I, Feger C, Galant C, et al. Mammography and subsequent whole-breast sonography of non-palpable breast cancers: the importance of radiologic breast density. *AJR Am J Roentgenol*. 2003;180(6):1675-1679.
- Michaelson JS, Silverstein M, Sgroi D, et al. The effect of tumor size and lymph node status on breast carcinoma lethality. *Cancer*. 2003;98(10):2133-2143.
- Michaelson JS, Silverstein M, Wyatt J, et al. Predicting the survival of patients with breast carcinoma using tumor size. *Cancer*. 2002;95(4):713-723.
- Berg WA. Supplemental screening sonography in dense breasts. *Radiol Clin North Am*. 2004;42(5):845-851.
- Bosch AM, Kessels AG, Beets GL, et al. Inter-examination variation of whole breast ultrasound. *Br J Radiol*. 2003;76(905):328-331.
- Berg WA, Blume J, Cormack JB, Mendelson EB. Operator dependence of physician-performed whole breast US: lesion detection and characterization. *Radiology*. 2006;241(2):355-366.
- Farria DM, Schmidt ME, Monsees BS, et al. Professional and economic factors affecting access to mammography: a crisis today, or tomorrow? *Cancer*. 2005;104(3):491-498.
- Corsetti V, Ferrari A, Ghirardi M, et al. Role of ultrasonography in detecting mammographically occult breast carcinoma in women with dense breasts. *Radiol Med (Torino)*. 2006;111(3):440-448.
- Saslow D, Boetes C, Burke W, et al. American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. *CA Cancer J Clin*. 2007;57(2):75-89.
- Berg WA, Blume JD, Cormack JB, Mendelson EB, Madsen EL. Lesion detection and characterization in a breast US phantom: results of the ACRIN 6666 Investigators. *Radiology*. 2006;239(3):693-702.
- Berg WA, Mendelson EB, Blume JD, Cormack JB. Performance of experienced breast imagers in interpreting breast ultrasound: results of training in the ACRIN 6666 protocol [abstract]. *AJR Am J Roentgenol*. 2004;182(4):34.
- Mendelson EB, Baum JK, Berg WA, Merritt CRB, Rubin E. *Breast Imaging Reporting and Data System, BI-RADS: Ultrasound*. Reston, VA: American College of Radiology; 2003.
- Rosenberg RD, Yankaskas BC, Abraham LA, et al. Performance benchmarks for screening mammography. *Radiology*. 2006;241(1):55-66.
- Leisenring W, Alonzo T, Pepe MS. Comparisons of predictive values of binary medical diagnostic tests for paired designs. *Biometrics*. 2000;56(2):345-351.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845.
- Metz CE, Wang P, Kronman HA. New approach for testing the significance of differences between ROC curves measured from correlated data. In: Deconinck F, ed. *Information Processing in Medical Imaging*. The Hague, the Netherlands: Nijhoff; 1984.
- Zhou X-H, Obuchowski NA, McClish DK. *Statistical Methods in Diagnostic Medicine*. New York, NY: John Wiley; 2002.
- Adams-Campbell LL, Makambi KH, Palmer JR, Rosenberg L. Diagnostic accuracy of the Gail model in the Black Women's Health Study. *Breast J*. 2007;13(4):332-336.
- Bondy ML, Newman LA. Breast cancer risk assessment models: applicability to African-American women. *Cancer*. 2003;97(1)(suppl):230-235.
- Gail MH, Costantino JP, Pee D, et al. Projecting individualized absolute invasive breast cancer risk in African American women. *J Natl Cancer Inst*. 2007;99(23):1782-1792.
- 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*. 2004;292(11):1317-1325.
- Kuhl CK, Schradang S, Leutner CC, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *J Clin Oncol*. 2005;23(33):8469-8476.
- Sardanelli F, Podo F, D'Agnolo G, et al. Multi-center comparative multimodality surveillance of women at genetic-familial high risk for breast cancer (HIBCRI study): interim results. *Radiology*. 2007;242(3):698-715.
- 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*. 2007;244(2):381-388.
- Menell JH, Morris EA, Dershaw DD, Abramson AF, Brogi E, Liberman L. Determination of the presence and extent of pure ductal carcinoma in situ by mammography and magnetic resonance imaging. *Breast J*. 2005;11(6):382-390.
- Schnall MD, Blume J, Bluemke DA, et al. Diagnostic architectural and dynamic features at breast MR imaging: multicenter study. *Radiology*. 2006;238(1):42-53.
- Berg WA. Beyond standard mammographic screening: mammography at age extremes, ultrasound, and MR imaging. *Radiol Clin North Am*. 2007;45(5):895-906, vii.
- Booi RC, Carson PL, O'Donnell M, Roubidoux MA, Hall AL, Rubin JM. Characterization of cysts using differential correlation coefficient values from two dimensional breast elastography: preliminary study. *Ultrasound Med Biol*. 2008;34(1):12-21.
- Kolb TM, Lichy J, Newhouse JH. Occult cancer in women with dense breasts: detection with screening US—diagnostic yield and tumor characteristics. *Radiology*. 1998;207(1):191-199.
- Schell MJ, Yankaskas BC, Ballard-Barbash R, et al. Evidence-based target recall rates for screening mammography. *Radiology*. 2007;243(3):681-689.
- Swan JS, Lawrence WF, Roy J. Process utility in breast biopsy. *Med Decis Making*. 2006;26(4):347-359.