Thomas M. Kolb, MD
Jacob Lichy, MD
Jeffrey H. Newhouse, MD

Index terms:
Breast, US, 00.129, 00.12989
Breast neoplasms, diagnosis, 00.30
Breast radiography, 00.11
Cancer screening, 00.11. 00.129, 00.12989

Published online before print
10.1148/radiol.2251011667
Radiology 2002; 225:165–175

Abbreviations:
BI-RADS = Breast Imaging Reporting and Data System
HRT = hormonal replacement therapy
PE = physical examination

[1] From 222 E 68th St, New York, NY 10021 (T.M.K., J.L.) and Department of Radiology, Columbia-Presbyterian Medical Center, New York, NY (T.M.K., J.H.N.). From the 1998 RSNA scientific assembly. Received October 11, 2001; revision requested December 5; final revision received April 11, 2002; accepted April 18. **Address correspondence to** T.M.K. (e-mail: *tkolb@panix.com*).

Author contributions:
Guarantor of integrity of entire study, T.M.K.; study concepts, T.M.K., J.L., J.H.N.; study design, T.M.K.; literature research, T.M.K.; clinical studies, T.M.K.; data acquisition, T.M.K.; data analysis/interpretation, T.M.K., J.H.N.; statistical analysis, T.M.K.; manuscript preparation, editing, and final version approval, T.M.K., J.H.N.; manuscript definition of intellectual content and revision/review, T.M.K., J.L., J.H.N.

© RSNA, 2002

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[1]

PURPOSE: To *(a)* determine the performance of screening mammography, ultrasonography (US), and physical examination (PE); *(b)* analyze the influence of age, hormonal status, and breast density; *(c)* compare the size and stage of tumors detected with each modality; and *(d)* determine which modality or combination of modalities optimize cancer detection.

MATERIALS AND METHODS: A total of 11,130 asymptomatic women underwent 27,825 screening sessions, (mammography and subsequent PE). Women with dense breasts subsequently underwent screening US. Abnormalities were deemed positive if biopsy findings revealed malignancy and negative if findings from biopsy or all screening examinations were negative.

RESULTS: In 221 women, 246 cancers were found. Sensitivity, specificity, negative and positive predictive values, and accuracy of mammography were 77.6%, 98.8%, 99.8%, 35.8%, and 98.6%, respectively; those of PE, 27.6%, 99.4%, 99.4%, 28.9%, and 98.8%, respectively; and those of US, 75.3%, 96.8%, 99.7%, 20.5%, and 96.6%, respectively. Screening breast US increased the number of women diagnosed with nonpalpable invasive cancers by 42% (30 of 71). Mammographic sensitivity declined significantly with increasing breast density ($P < .01$) (48% for the densest breasts) and in younger women with dense breasts ($P = .02$); the effects were independent. Mammography and US together had significantly higher sensitivity (97%) than did mammography and PE together (74%) ($P < .001$). Tumors detected at mammography and/or US were significantly smaller ($P = .01$) and of lower stage ($P = .01$) than those detected at PE.

CONCLUSION: Mammographic sensitivity for breast cancer declines significantly with increasing breast density and is independently higher in older women with dense breasts. Addition of screening US significantly increases detection of small cancers and depicts significantly more cancers and at smaller size and lower stage than does PE, which detects independently extremely few cancers. Hormonal status has no significant effect on effectiveness of screening independent of breast density.
© RSNA, 2002

Mammography and palpation are the currently accepted breast cancer screening tests. Their effectiveness is imperfectly known due to differences among the reported series, less-than-ideal standards for defining true-negative and false-negative examination findings, lack of analysis of patient subgroups, and variation in risk factors and characteristics of the normal breast tissues.

Mammography has been evaluated in seven large randomized controlled trials and in smaller nonrandomized studies (1–8). Detection rates vary widely: Sensitivities have been reported from 68% to 88% and specificities from 82% to 98%. Sensitivity is higher in women aged 50 years and older than in those aged 40–49 years old, and there is a larger mortality benefit in older women (7–16). Sensitivity is lower in radiographically dense breasts (17–20), which are often seen in young premenopausal women or in those who receive hormonal replacement therapy (HRT). But mammographic performance as a function of age, hormonal status, and breast density has been incompletely evaluated due to inconsistent density measurements, other confounding factors, and reliance on interval cancers to calculate sensitivity.

Authors of three nonrandomized studies (21–23), in which the independent effects of age, breast density, and hormonal status on mammographic performance were examined, have reported conflicting results. Furthermore, the sensitivity of mammography for each of the American College of Radiology Breast Imaging Reporting and Data System (BI-RADS) (24) density categories has never been reported, to our knowledge.

Screening physical examination (PE) performed by a physician, although widely used and recommended by a number of medical organizations (25–27), has not been as rigorously tested and evaluated as mammography, nor has a mortality benefit been demonstrated. The performance of screening PE related to patient age has been reported in only two studies (28, 29); in neither of the studies was its performance in any other subpopulation of women evaluated, and no analysis of tumor size and stage has been reported, to our knowledge.

The ability of screening ultrasonography (US) to depict nonpalpable mammographically occult cancers at a size and stage similar to those detected at mammography has recently been reported (30). Cancers identified by using screening US are not unimportant, especially since they usually occur in women with dense breasts who are most often young. Comparison of the performance of screening breast US with that of screening mammography and PE can help identify the contributions of each modality toward breast cancer detection.

We report the results of a prospective study in which mammography, PE, and US were used to screen a large population of asymptomatic woman. Our purposes were *(a)* to determine the performance of screening mammography, US, and PE; *(b)* to analyze the influence of age, hormonal status, and breast density; *(c)* to compare the size and stage of tumors detected with each modality; and *(d)* to determine which modality or combinations of screening techniques optimize cancer detection.

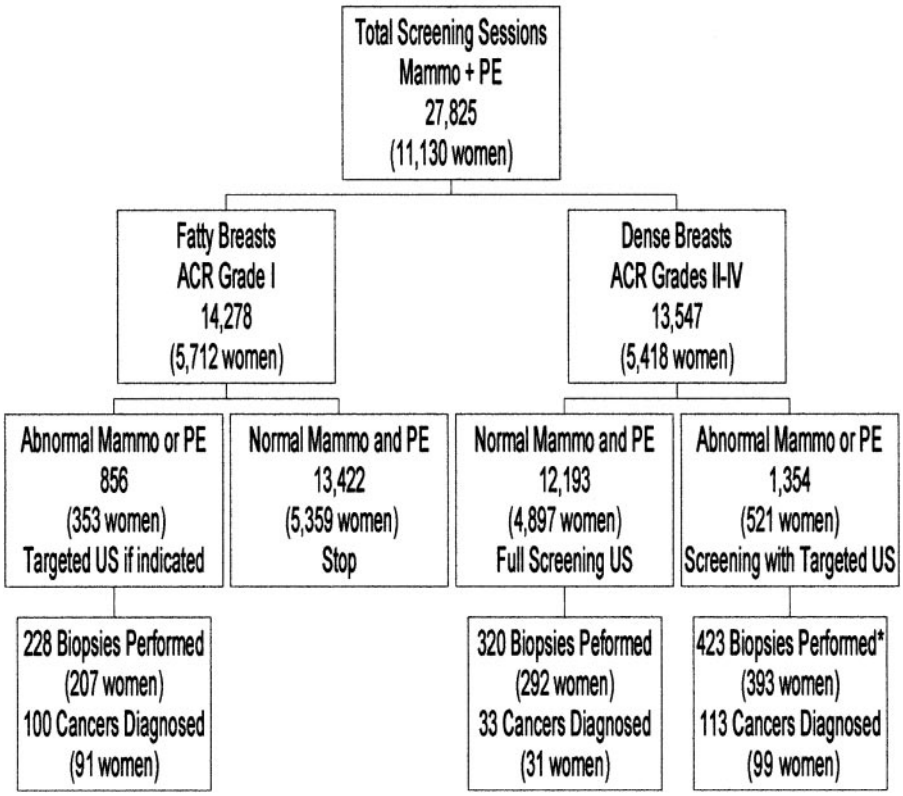

Figure 1. Flow diagram for patients examined shows the screening detection scheme. In 31 women with a known mammographic *(Mammo)* and/or physical finding, biopsy was performed for 38 masses (*) detected at US alone. Four cancers were diagnosed in three women. ACR = American College of Radiology BI-RADS.

MATERIALS AND METHODS

Between January 15, 1995, and September 30, 2000, 27,825 screening sessions, which included PE and mammography, were performed in 11,130 women (mean age, 59.6 years; SD, 15.8) who had no signs or symptoms of breast abnormalities (Figure). In this series, 14,278 examinations were performed in 5,712 women (mean age, 63.7 years; SD, 13.8) with fatty breasts. The 13,547 examinations performed in 5,418 women (mean age, 54.7 years; SD, 15.3) with dense breasts (BI-RADS density categories 2–4) were followed by bilateral whole-breast screening US. A total of 69,197 screening examinations were performed in the study group.

Breast density was graded by a single examiner (T.M.K.) according to the following BI-RADS density categories (24): category 4, "extremely dense breast, which lowers the sensitivity of mammography"; category 3, "breast tissue is heterogeneously dense, which may lower the sensitivity of mammography"; category 2, "scattered fibroglandular densities that could obscure a lesion on a mammogram"; and category 1, "breast [composed] almost entirely [of] fat." For the purpose of this study, breast density categories 2, 3, and 4 were classified as "dense;" and breast density category 1, as "fatty." Our method for categorizing breast density has been previously described (30).

Order of Screening Examinations

Each woman was first evaluated with routine screening mammography and then with a complete PE of the breasts. The first 700 women with fatty breasts, defined as category 1, also underwent a bilateral screening breast US. No additional cancers were found with US alone.

After it was determined that screening breast US in women with fatty breasts was sufficiently unlikely to depict otherwise occult cancers, the remainder of the women with fatty breasts were not screened with US. Therefore, the performance characteristics of screening US in women with fatty breasts were not determined, and the false-negative rate of screening mammography was calculated on the basis of the results of the concurrent PE alone.

In all women with dense breasts (BI-RADS density categories 2–4), a complete bilateral screening US was performed and we defined the performance characteristics of screening US for only those women. If there was a mammographically or physically detected mass and/or a mammographically detected grouping of suspicious calcifications without a mammographically associated mass but within a background of dense glandular tissue, then, in addition to appropriate mammographic spot compression and/or magnification views, the area questioned was interrogated during the course of the screening US.

We adhered to the guidelines contained in the Declaration of Helsinki principles (31) when we performed screening breast US. Each patient was given a complete explanation of the US procedure by the examining physician (T.M.K.) and was informed that if a mass was detected, depending on its appearance, she would need either follow-up US examination, US-guided percutaneous biopsy, or surgical biopsy. As the study progressed, data were analyzed and our earliest findings were published (30). Patients were thereafter given more specific cancer detection and false-positive rates for screening US. Verbal consent was obtained from each patient.

Performance of Screening Examinations

All evaluations were performed by a single radiologist (T.M.K.) who specializes in breast cancer detection. The examiner knew the results of each preceding examination prior to performing the succeeding examination; that is, the results of mammography were known at the time of the PE, and the results of both were known at the time of screening US. Each patient was interviewed to determine age, personal and family history of breast cancer, menopausal status, and the use of exogenous hormones. Menopause was defined as no menses for 12 months prior to the screening examination. HRT was defined as application or ingestion of any estrogen-related compound. Women who had undergone hysterectomy but not oophorectomy were arbitrarily placed either in the postmenopausal group, if presenting for screening at age 50 years and older, or in the premenopausal group, if 49 years or younger.

All mammographic examinations were performed with dedicated mammography machines (DMR [1995 to present], 800T [1997 to present], and 500T [1995–1996]; GE Medical Systems, Milwaukee, Wis) by using a 0.3-mm focal spot and 0.1 mm for magnification. Dedicated mammography cassettes (Min R-2; Kodak, Rochester, NY) and screens (Min-R; Kodak) were used. Film processing optimized for the mammographic units was used. Each mammographic examination was monitored for optimal exposure, contrast, and positioning at the time of processing during each patient's visit. Eighty-four percent of the time, prior mammograms were available for comparison with those of the current study.

All US examinations were performed with commercially available digital US machines (HDI 5000 [1997–2000] and HDI Ultramark 9 with ESP [1995–1997]; Advanced Technology Laboratories, Bothell, Wash) by using a broadband linear probe (L12-5 [HDI 5000] or L10-5 [Ultramark 9]; Advanced Technology Laboratories) with a 3.8-cm-wide field of view at a 4-cm depth of view.

All US examinations included anatomic regions that covered the entire breast. The field of view was adjusted for breast size. Focal zones were limited (usually one to two were used) to ensure a high enough frame rate to facilitate screening. The focal zone(s) was initially placed in the far field or near the level of the pectoral muscle. Patients lay supine with arms stretched over their heads. They were commonly shifted to the contralateral posterior oblique position to facilitate scanning the lateral and inferior aspects of the breast. Patients with larger breasts often needed to be turned to the ipsilateral posterior oblique position to scan the medial portion of the breast. In most women, repositioning from the supine position was necessary.

Higher contrast settings were used to facilitate the search for subtle architectural changes and isoechoic masses. If a lesion was detected, targeted analysis was begun by decreasing the field of view and realigning the focal zone (if appropriate) to the level of the suspected abnormality. Images were obtained at varied angles by turning the transducer, commonly rotating it 360° through the area in question, thus changing the transducer-to-lesion angle. Patients' position was often shifted to change the skin-to-lesion angle, as well. Finally, if the findings from screening examination were negative, a single hard-copy image of a normal area of a single breast was obtained to document the performance of the examination.

We have reported (30) that the average time to perform screening breast US in women with normal mammographic and PE findings was 3 minutes 59 seconds (range, 1 minute 28 seconds to 9 minutes 46 seconds). In this study, examinations were timed in 50 patients with dense breasts and mammographic and/or palpable findings in whom screening US was performed with additional attention to the site(s) of concern. The mean time to perform the entire US evaluation was 4 minutes 39 seconds (range, 2 minutes 11 seconds to 11 minutes 30 seconds) and depended on the size of the breasts and the pathologic findings.

PE consisted of visual inspection in the erect and supine positions (upright only for women with small fatty breasts), palpation of the entire breast in the supine position (upright for women with small fatty breasts), with arms overhead and the patient turned to the contra- and ipsilateral oblique positions, as necessary. The finger pads and a rotational movement at each point of palpation with a spoke search pattern were used. The axillae and supraclavicular triangles were also palpated.

Findings from PE were considered abnormal if a discrete mass, area of asymmetric thickening, bloody or serous discharge, skin retraction, or nipple or areolar rash were identified.

All patients scheduled for surgical biopsy on the basis of a recommendation from results of a screening examination or percutaneous biopsy based on screening results were reexamined by a surgeon specializing in breast surgery. The surgeons had available to them mammograms, US images, and written reports detailing the exact location and the degree of suspicion of the abnormality. The definition of nonpalpability in these instances was whether a needle localization procedure was necessary to guide the surgeon for surgical excision.

The majority of women (78%) had also undergone clinical breast examination performed by their primary physicians, gynecologists, or breast surgeons within 1 month prior to the study examinations. The remaining 22% of women either underwent a clinical breast examination more than 1 month prior to the screen-

ing examination or were to undergo one after the mammographic results were available. Only women with no signs or symptoms of a breast abnormality were admitted to this study.

The mammographic appearance of cancer in different subpopulations was recorded. Since mammographic sensitivities are different among some populations, if there were differences in the appearances of tumors among populations this analysis might determine whether specific appearances might be associated with different mammographic sensitivities. Mammographic findings were divided to two groups: those that were calcified and/or deformed the normal breast architecture, which we speculated might be detectable in both fatty and dense breasts, and those that were noncalcified and nondeforming, which we theorized might be more difficult to detect in dense breasts. The former category included isolated calcifications with no associated mass, calcifications associated with any mass, masses that were irregularly marginated or spiculated, or isolated architectural distortion. The latter category included isolated patchy or asymmetric densities and well-circumscribed masses.

Stereotactic, core, and vacuum-assisted biopsy; US-guided core biopsy; fine-needle aspiration; and surgical excision were performed to establish the diagnoses for masses and calcifications, which were suspected to indicate malignancy. If at PE a clinically suspicious mass was detected, even if imaging findings were negative, either percutaneous biopsy was performed or surgical biopsy was recommended. All mammographically identified masses were interrogated with US; the decision to perform biopsy was made on the basis of both mammographic and US findings. Masses considered for biopsy based on US findings had irregular margins and/or shape.

Biopsy was performed on mammographically detected masses that were not sonographically visualized if they were irregularly marginated or spiculated, if they had not been previously documented, if they had changed to a more malignant appearance, or at the patient's request. Biopsy was performed on mammographically identified calcifications if they had any malignant appearance, such as pleomorphism or a branched ductal pattern. To assess for the presence of an associated mass, US was used to interrogate areas of mammographically suspicious calcifications if the calcifications were within a background of fibroglandular tissue judged dense enough to be able to obscure a lesion; if no mass was found, however, US was never used to determine the likelihood that the calcifications represented malignancy. Masses for which a patient or the patient's clinician insisted on biopsy, even without a radiologic recommendation, were also examined at biopsy and reported here as such.

Not all solid masses were subjected to biopsy. During US, patients were informed of the presence of any masses and the likelihood that they represented malignancies. On the basis of the judgment of the examiner, masses were either subjected to biopsy, placed in a follow-up category, or continued to be screened annually. If after adequate percutaneous biopsy the results were positive, suspicious, discordant, or inconclusive with a strong imaging suspicion for cancer, the patient was referred for surgical excision. Results of 53 lesions for which biopsy was recommended were unavailable for review either because the patients refused biopsy and/or were lost to follow-up.

Follow-up

All patients with benign percutaneous biopsy results were followed up 4–6 months and 1 year later to confirm benignity. In some cases, the patients' clinicians requested surgical biopsies regardless of radiologic recommendation. None of these biopsy findings proved to be malignant.

Repeat examinations of all mammographic and US findings believed to be benign but placed in a follow-up category were performed at 6 months, 1 year, and 2 years, with the modality that was best able to help evaluate the finding. Patients returning for follow-up were not considered as asymptomatic screening patients. Cancers in the same breast were considered separate and multiple if they were at least 1 cm apart.

Risk Factors

A total of 16.7% of women in the patient population had a personal history of breast cancer, 23.9% had any family history, and 16.2% had a primary family (mother, sister, daughter, father or brother) history of breast cancer. For the purpose of this study, women with a personal history or with a primary family member with a history of breast cancer and those with a previous high-risk biopsy results were considered to be at high risk for eventually developing breast cancer. Women without these risk factors were considered at normal risk. Thirty percent of all examined women and 26.5% of those with dense breasts were at high risk for subsequently developing breast cancer. The prevalence of cancers detected with each modality in both high- and normal-risk groups was recorded and compared.

The sizes of the cancer excised and measured at pathologic examination were available in 223 of the 246 cancers detected. One hundred eight-six of the 221 women with a diagnosis of cancer underwent staging. For 18 women, staging data were unavailable, and 17 women had recurrent cancer in the ipsilateral breast. For staging, five stage groupings (stage 0–IV) in the TNM classification were used (32). Patients with recurrent cancer in the ipsilateral breast, primary breast lymphoma, or sarcoma did not undergo staging.

Statistical Analyses

The total number of biopsies performed and cancers detected was assessed for the total patient population, as well as for subpopulations grouped by age (50 years and older and 49 years and younger), hormonal status (premenopausal, postmenopausal not receiving HRT, and postmenopausal receiving HRT), and breast density (BI-RADS categories 1–4). Biopsy and cancer detection rates were reported for each modality, both for lesions detected solely with that modality and for lesions detected with a combination of that modality plus one or both of the other two screening examinations.

Performance characteristics of each screening modality, including sensitivity, specificity, positive and negative predictive values, and accuracy, were calculated. Sensitivity was defined as the percentage of cancers detected (with a specific modality) among all cancers detected with any modality: TP/(TP + FN), where TP is true-positive and FN is false-negative. Specificity was defined as the percentage of normal results from examination (with a specific modality) of any area of the breast where cancer was not detected with any modality: TN/(TN + FP), where TN is true-negative and FP is false-positive. The positive predictive value was defined as the percentage of cancers detected (with a specific modality) among those lesions for which the modality yielded positive results: TP/(TP + FP). The negative predictive value was defined as the proportion of normal results from examination (with a specific modality) of areas where cancer was not detected with

TABLE 1
Performance Characteristics of Each Modality

Modality	Sensitivity (%)	Specificity (%)	NPV (%)	PPV (%)	Accuracy (%)
Mammography	77.6 (191/246)	98.8 (27,237/27,579)	99.8 (27,237/27,292)	35.8 (191/533)	98.6 (27,428/27,825)
PE	27.6 (68/246)	99.4 (27,412/27,579)	99.4 (27,412/27,590)	28.9 (68/235)	98.8 (27,480/27,825)
US	75.3 (110/146)	96.8 (12,975/13,401)	99.7 (12,975/13,011)	20.5 (110/536)	96.6 (13,085/13,547)

Note.—Screening US was performed and its results reported only in women with dense breasts. Calculations include both invasive and noninvasive cancers. Data in parentheses are numbers used to calculate the percentages. See "Statistical Analyses" in Materials and Methods for definitions of diagnostic performance statistics. NPV = negative predictive value, PPV = positive predictive value.

any other modality among all areas where the specific modality yielded normal results: TN/(TN + FN). Accuracy was defined as the percentage of lesions in which a specific modality correctly predicted the presence or absence of cancer, among all examinations performed: (TP + TN)/(TP + TN + FP + FN).

An examination finding revealing a visible or palpable lesion was classified as true-positive if biopsy findings revealed cancer and false-positive if biopsy or follow-up findings revealed no cancer. An examination finding revealing no abnormality was classified as true-negative if all other examination findings were normal and false-negative if any other examination findings were abnormal and biopsy revealed cancer. As detailed earlier, a single examiner performed each of the screening studies and knew the results of the preceding studies before performing the next one. Therefore, the sensitivity of PE presumes the examiner knew the mammographic result, and the sensitivity of US presumes the examiner knew both the mammographic and the PE results.

Sensitivities calculated in such a manner for each modality and groups of women identified by age, hormonal status, and breast density were computed. The sensitivities of each modality within each subpopulation were compared by using the Fisher exact test. Multivariate logistic regression analysis was used to assess the individual contributions of age, density, and hormonal status on mammographic sensitivity. The mean size and stage of cancers detected with each modality were compared by using the Student *t* and the Wilcoxon rank sum tests. Morphologic characteristics of mammographically identified cancers grouped by age and density are reported and compared by using the Fisher exact test. The sensitivity of, as well as the size and stage of tumors detected with, conventional screening (mammography plus PE) were compared with those that would have been detected with mammography and US if palpation had not been performed by using the Fisher exact test and the Wilcoxon rank test. A *P* value of less than .05 was used to indicate significant differences between groups.

RESULTS

Biopsies and Cancers Detected

Total population.—After 27,825 screening sessions, 971 biopsies were performed in 892 women (3.2% biopsy rate for all women), and 246 cancers (25.3% positive biopsy rate) were diagnosed in 221 women (mean age, 59.9 years; SD, 14.5; 7.94 women with cancer per 1,000 screened women), including 42 (17.1%) foci of pure ductal carcinoma in situ and 204 (82.9%) invasive cancers.

Conventionally screened population.—In 27,825 women, 613 biopsies were performed on the basis of mammographic and/or PE findings (2.2% biopsy rate for women conventionally screened) for which US was used, when appropriate, to further characterize the suspected abnormality. Two hundred nine cancers, 41 pure ductal carcinomas in situ and 168 invasive, were diagnosed (34.1% positive biopsy rate with conventional screening). Of the 209 cancers, 18 (8.6%) were mammographically occult but were identified at PE.

Performance characteristics.—The overall performance characteristics of each screening modality for all cancers, invasive and noninvasive and whether or not detected with any other modality, are reported in Table 1. For invasive cancers only, the sensitivities of mammography, US (in women with dense breasts), and PE were 74.0% (151 of 204), 73.0% (92 of 126), and 30.4% (62 of 204), respectively.

Cancers detected with mammography, whether or not detected by any other means, had a mean size of 13.5 mm; in 81% (117 of 145) of women, they were stage 0 or I. Cancers detected with US, whether or not detected with any other technique, had a mean size of 14.7 mm; in 61% (88 of 145) of women, they were stage 0 or I. Cancers found at PE, whether or not detected with any other technique, had a mean size of 21.6 mm; in 14% (20 of 145) of women, they were stage 0 or I.

Nonpalpable cancers (n = 178) were smaller (mean size, 10.2 mm; SD, 6.0; $P <$ .01, Wilcoxon rank sum) and of lower stage (in 91.9% [125 of 136] women they were stage 0 or I; $P <$ .01, Fisher exact test) than were palpable cancers (n = 68; mean size, 21.6 mm; SD, 10.1; in 40% [20 of 50] women they were stage 0 or I).

A total of 334 biopsies were performed owing to findings with more than a single screening modality (34.4% of all biopsies), and 132 cancers were diagnosed (39.5% positive biopsy rate).

Biopsies and Cancers Detected on the Basis of Findings from Only One Screening Modality

Mammography.—A total of 263 biopsies (27.1% [263 of 971]) were performed owing solely to mammographic findings; findings from the other two screening modalities were unremarkable. Seventy-five cancers (30% of all cancers; 28.5% positive biopsy rate for findings identified only with mammography) in 65 women (29% of all women with cancer) were detected in this group; 36 cancers were invasive and 39 were noninvasive. Forty-five of the 75 cancers were detected in fatty breasts, and 30 were detected in dense breasts. In fatty breasts, mammography alone depicted 45% (45 of 100) of all cancers and 30% (24 of 79) of all invasive cancers. In dense breasts, mammography alone depicted 20.5% (30 of 146) of all cancers and 9.4% (12 of 127) of all invasive cancers. The mean size of all cancers detected only with mammography was 8.9 mm (SD, 5.7); in 73% (36 of 49) of women with sized cancers, they were 1 cm or smaller. In 98% of women, the cancers detected only with mammography were stage 0 or I.

PE.—Sixteen biopsies (1.6% of all biop-

sies) were performed solely on the basis of PE findings, findings from the other two modalities were unremarkable. In six (2.7%) of 221 women, six cancers (2.4% of all cancers; mean size, 21.8 cm; SD, 4.03)—all invasive, two in fatty breasts and four in dense breasts—were detected only at PE. Of the two cancers in fatty breasts not evident at mammography, one was in a woman with contracted subglandular implants and one was located high in the upper inner quadrant of the breast, which was not visualized at mammography. Both of these cancers were correctly identified on targeted US scans as malignant-appearing masses.

In dense breasts, palpation alone enabled detection of 2.7% (four of 146) of all cancers and 3.2% (four of 125) of all invasive cancers. The mean size of cancers detected only at PE was 21.8 mm (SD, 4.0); no cancers (zero of six) were 1 cm or smaller. In no women with cancers detected only at PE (zero of six) were the cancers detected at stage 0 or I.

Screening US.—In 13,547 women with dense breasts, 799 solid masses were detected with screening US alone. Of these, 441 (55.2%) were placed either in a follow-up category (n = 400, 50.0%) or were allowed to continue with annual screening (n = 41, 5.1%), none were determined to be cancer. A total of 358 biopsies (36.9% of all biopsies) in 13,547 women (2.6% biopsy rate) were performed on the basis of a finding from the screening US alone. Thirty-seven cancers (10.3% positive biopsy rate, 15% of all cancers)—36 invasive and one noninvasive—in 34 women (15.4% of all women with cancer) were detected only with screening US.

In women with dense breasts, US alone depicted 25.5% (37 of 145) of all cancers, 29.0% (36 of 124) of all invasive cancers, and 37.0% (37 of 100) of all nonpalpable cancers. US alone enabled detection of nonpalpable invasive cancer in 42% (30 of 71) of women in whom no other cancers were detected with any other screening modality and enabled detection of any type cancer and no other cancers detected with any other screening modality in 38% (13 of 34) of women younger than 50 years and 19% (18 of 96) of women age 50 years and older. The mean size of cancers detected with US only was 9.9 mm (SD, 6.79); 70% (26 of 37) of cancers were subcentimeter. In 89% (25 of 28) of women, the cancers detected only at US were stage 0 or I. Cancers detected at US alone were not significantly different in size and stage than those detected only at mammography.

The overall false-positive rate of screening US for masses requiring biopsy was 2.4% (321 of 13,547) and for those requiring biopsy or follow-up was 5.3% (762 of 13,547).

TABLE 2
Sensitivity of Each Modality for Cancer Detection in Women with Different Breast Densities

Modality	BI-RADS Category				
	1	2	3	4	2–4
Mammography	98.0 (98/100)	82.9 (34/41)	64.4 (38/59)	47.8 (22/46)	64.4 (94/146)
US	NP	65.9 (27/41)	81.4 (48/59)	76.1 (35/46)	75.3 (110/146)
PE	22.0 (22/100)	31.7 (13/41)	28.8 (17/59)	34.8 (16/46)	31.5 (46/146)

Note.—Data are percentages. Data in parentheses are numbers used to calculate percentages. NP = not performed.

TABLE 3
Sensitivity of Screening Modalities according to Age

Modality	Percentage of Women	
	49 Years or Younger	50 Years or Older
Mammography*	58.0 (29/50)	82.7 (162/196)
PE*	36.0 (18/50)	25.5 (50/196)
US†	78.6 (33/42)	74.0 (77/104)

Note.—Data in parentheses are numbers used to calculate the percentages.
* Women with both fatty and dense breasts.
† Only women with dense breasts (BI-RADS category 2–4).

Prevalence of Cancers Detected Only at Screening US

Among women with dense breasts, the prevalence of those with cancers detected only at screening US was 0.23% (31 of 13,547 women), with rates of 0.11% (three of 2,732 women), 0.27% (13 of 4,815 women), and 0.25% (15 of 6,000 women) in breast density categories 2, 3, and 4, respectively. Among high-risk women with dense breasts, the cancer prevalence rate was 0.42%, with rates of 0.15% (one of 674 women), 0.44% (six of 1,365 women), and 0.52% (eight of 1,549 women) in breast density categories 2, 3, and 4, respectively. There was a significant increase in the prevalence rate of cancers detected only at US in women with dense breasts who were at high risk versus women at normal risk (P = .012, Fisher exact test).

The addition of screening US in women with dense breasts increased the rate of negative biopsy findings from 65.9% in women conventionally screened with mammography and PE to 74.6% in women screened with all three modalities.

Biopsies and Cancer Detection in Women with Differing Breast Densities

In 14,278 women (mean age, 63.6 years; SD, 13.7) with fatty breasts, 228 (1.6% biopsy rate) biopsies were performed. In 91 women, 100 cancers (mean size, 12.7 mm; SD, 8.56)—79 invasive and 21 noninvasive—were diagnosed (43.8% positive biopsy rate, 84% [62 of 74] stage 0 or I; 91/14,278 = 6.4 cancers per 1,000 women screened).

In 13,547 women (mean age, 54.7 years; SD, 15.2) with dense breasts, 743 (5.5% biopsy rate) biopsies were performed. In 130 women, 146 (19.7%) cancers (mean size, 14.2 mm; SD 9.3)—125 invasive and 21 noninvasive—were diagnosed (74.1% [83 of 112] stage 0 or I; 130/13,547 = 9.6 cancers per 1,000 women screened). Of the 743 biopsies, 385 were performed on the basis of mammographic or PE findings (2.8% biopsy rate for conventional screening), and 109 (28.3%) cancers were detected.

Both the biopsy ($P < .001$) and the cancer detection rates per 1,000 women screened (P = .003) were higher in women with dense breasts, who were evaluated either conventionally or with the addition of screening US, than in those with fatty breasts (Fisher exact test).

There was a significant decrease in the mammographic sensitivity for all cancers from category 1 (98%) through category

TABLE 4
Sensitivity of Modalities in Women with Differing Hormonal Status

Modality	Percentage of Women		
	Premenopausal	Postmenopausal not Receiving HRT	Postmenopausal Receiving HRT
Mammography*	66.7 (40/60)	80.4 (119/148)	84.2 (32/38)
PE*	38.3 (23/60)	25.0 (37/148)	21.0 (8/38)
US†	71.4 (35/49)	82.2 (60/73)	62.5 (15/24)

Note.—Data in parentheses are numbers used to calculate the percentages. HRT = hormonal replacement therapy.
* Women with both fatty and dense breasts.
† Only women with dense breasts (BI-RADS category 2–4).

TABLE 5
Combined Effect of Age and Breast Density on Mammographic Sensitivity

Age	BI-RADS Category					*P* Value*
	1	2	3	4	2–4	
<50 y	100 (8/8)	100 (3/3)	50.0 (10/20)	42.1 (8/19)	50.0 (21/42)	<.015
≥50 y	97.8 (90/92)	81.6 (31/38)	71.8 (28/39)	51.9 (14/27)	70.2 (73/104)	<.001
P value*	>.99	>.99	.15	.56	<.035	

Note.—Data are percentages. Data in parentheses are numbers used to calculate the percentages.
* BI-RADS category 1 versus 2–4, Fisher exact test ($P < .05$).

4 (47.8%) breast density ($P < .001$, χ^2 test) and between category 1 (98%) and all dense breast categories (categories 2–4) combined (64.4%; $P < .001$, Fisher exact test) (Table 2). The mammographic sensitivity for invasive cancer in category 4 breast density was 44% (18 of 41). There were no significant changes in the sensitivity of PE in women with both fatty and dense breasts or in the sensitivity of US in women with dense breasts.

Biopsies and Cancer Detection in Women of Differing Ages

In 5,826 women younger than 50 years (mean age, 41.9 years; SD, 9.1), 352 (6.0%) biopsies were performed, and 50 cancers (mean size, 13.6 mm; SD, 7.4)—45 invasive and five noninvasive—in 42 women (64% [21 of 33]) were stage 0 or I) were diagnosed (42/5,826 = 7.2 cancers per 1,000 women screened).

In 21,999 women 50 years and older (mean age, 64.1 years; SD, 14.1), 619 (2.8%) biopsies were performed, and 196 cancers (mean size, 13.6 mm; SD, 9.5)—159 invasive and 37 noninvasive—in 179 women (81.0% [124 of 153] stage 0 or I) were diagnosed (179/21,999 = 8.1 cancers per 1,000 women screened).

The biopsy rate, evaluated either conventionally or with the addition of screening US, was higher in women aged 49 years and younger than in older women ($P < .001$, Fisher exact test). However, there was no significant difference in the cancer detection in these two groups.

There was a significant decrease in mammographic sensitivity in women aged 49 years and younger compared with those 50 years and older ($P < .001$, Fisher exact test) (Table 3). Among the different age groups, there was no significant difference in the sensitivity of PE in women with both fatty and dense breasts nor that of US in women with dense breasts.

Biopsies and Cancer Detection in Women of Differing Hormonal Status

In 5,318 premenopausal women (mean age, 47.6 years; SD, 9.4), 246 (4.6% biopsy rate) biopsies were performed, and 60 cancers—50 invasive and 10 noninvasive (mean size, 13.4 mm; SD, 7.15)—in 51 women (58% [22 of 38] stage 0 or I) were diagnosed (51/5,318 = 9.6 cancers per 1,000 women screened).

In 17,765 postmenopausal women (mean age, 64.5 years; SD, 15.9) not receiving HRT, 637 (3.6%) biopsies were performed, and 148 cancers (mean size, 14.1 mm; SD, 9.50)—125 invasive and 23 noninvasive—in 135 women (80% [92 of 115] stage 0 or I) were diagnosed (135/17,765 = 7.6 cancers per 1,000 women screened).

In 4,742 postmenopausal women (mean age, 60.3 years; SD, 13.8) receiving HRT, 88 (1.9% biopsy rate) biopsies were performed, and 38 cancers (mean size, 12.3 mm; SD, 10.05)—29 invasive and nine noninvasive—in 35 women (94% [31 of 33] stages 0 or 1) were diagnosed (35/4,742 = 7.4 cancers per 1,000 women screened).

The biopsy rate, evaluated either conventionally or with the addition of screening US, was higher in premenopausal women than in postmenopausal women who were either receiving HRT ($P < .001$), not receiving HRT ($P < .001$), or among all postmenopausal women combined ($P < .001$, Fisher exact test). The rate was also higher in postmenopausal women not receiving HRT versus those that were receiving HRT ($P < .001$). But there was no difference in postmenopausal women not receiving HRT versus a combination of premenopausal women and postmenopausal women receiving HRT. There was no significant difference in the sensitivity of mammography, PE (in women with both fatty and dense breasts), or US (in women with dense breasts) in the detection of cancer among women of differing hormonal status (Table 4).

Effect of Density, Age, and Hormonal Status on Mammographic Sensitivity

Both in the younger and the older group, breast density at mammography was significantly inversely related to mammographic sensitivity (Table 5). Among women with dense breasts (BI-RADS category 2–4), mammography was significantly more sensitive ($P = .03$, Fisher exact test) in older women.

Of the 21 cancers that mammography did not depict in women with dense breasts and 49 years and younger, 20 (95.2%) cancers were depicted with screening US (mean size, 11.1 mm; SD, 5.7) and four (19%) were detected at PE (mean size, 19.5 mm; SD, 2.5).

The combined effect of hormonal status and breast density on mammographic sensitivity is described in Table 6. If the women in the three hormonal status groups are analyzed separately, in each group, mammographic sensitivity is significantly lower in women with dense breasts than in those with fatty breasts. However, in women with breasts of similar density but differing hormonal status, there is no significant change in mammographic sensitivity.

Of the 20 cancers that mammography

TABLE 6
Combined Effect of Hormonal Status and Breast Density on Mammographic Sensitivity

Hormonal Status	BI-RADS Category 1	2	3	4	2–4	*P* Value*
Premenopausal	100 (11/11)	100 (8/8)	58.8 (10/17)	45.8 (11/24)	59.2 (29/49)	<.01
Postmenopausal receiving HRT	100 (14/14)	71.4 (5/7)	92.3 (12/13)	25.0 (1/4)	75.0 (18/24)	<.042
Postmenopausal not receiving HRT	96.0 (72/75)	80.8 (21/26)	55.2 (16/29)	55.6 (10/18)	64.4 (47/73)	<.001
P value*	.60	.30	.06	.46	.42	

Note.—Data are percentages. Data in parentheses are numbers used to calculate the percentages. HRT = hormonal replacement therapy.
* BI-RADS category 1 versus 2–4, χ^2 ($P < .05$).

did not depict in premenopausal women with dense breasts, 18 (90.0%) were depicted with screening US (mean size, 10.5 mm; SD, 5.8) and four (20%) were detected at PE (mean size, 18.5 mm; SD, 3).

Multivariate logistic regression analysis was used to assess the individual contributions of age, density, and hormonal status on the sensitivity of mammography and PE. The analysis was limited by the high correlation between age, density, and hormonal status; however, there was a significant increase in the sensitivity for cancer detection in women with category 1 breast density versus those with categories 2–4 breast density and to a lesser, but significant, degree in older versus younger women. The sensitivity of PE was not significantly different among the subpopulations analyzed (Table 7).

TABLE 7
Multivariate Logistic Regression Analysis of Contributions of Age, Density, and Hormonal Status on Sensitivity of Mammography and PE

Variable	Sensitivity of Mammography *P* Value	Sensitivity of Mammography Odds Ratio	Sensitivity of PE *P* Value	Sensitivity of PE Odds Ratio
Hormonal status*	.364	0.624 (0.225–1.726)	.17	0.545 (0.227–1.308)
Breast density†	<.001	0.063 (0.019–0.211)	.25	1.435 (0.777–2.650)
Age‡	.032	0.324 (0.115–0.910)	.90	0.939 (0.368–2.396)

Note.—Screening US was not performed in women with category 1 breast density; therefore, its overall sensitivity in the subgroups analyzed cannot be reported. Data in parentheses are CIs.
* Grouped as either premenopausal or postmenopausal and receiving hormonal replacement therapy versus postmenopausal and not receiving hormonal replacement therapy.
† Grouped as BI-RADS category 1 versus categories 2–4.
‡ Grouped as 49 years and younger versus 50 years and older.

TABLE 8
Sensitivity of Combined Detection Modalities for Each Breast Density Category

Modality	BI-RADS Category 1	2	3	4	2–4
Mammography, PE, or both	100 (100/100)	87.8 (36/41)	74.6 (44/59)	63.0 (29/46)	74.7 (109/146)
Mammography, US, or both*	NP	100 (41/41)	98.3 (58/59)	93.5 (43/46)	97.3 (142/146)

Note.—Data are percentages. Data in parentheses are numbers used to calculate the percentages. NP = not performed.
* Screening US was not performed in women with BI-RADS category 1 breast density.

Morphologic Characteristics of Mammographic Findings in Women of Differing Ages and Breast Densities

There was no significant difference in the mammographic appearance of detected cancers (calcified and/or deforming the normal breast architecture vs noncalcified and nondeforming) in women 49 years and younger (21 [72%] of 29) versus that in women 50 years and older (126 [77.8%] of 162). When the morphologic analysis was restricted to women with only BI-RADS density category 3 and 4 breasts, while the differences between the two groups increased, again no significant difference was found between the two morphologic appearances in the younger age group (17 [94%] of 18 cancers) versus those in the older age group (32 [76%] of 42 cancers).

Effect of Substituting US for PE in Addition to Mammography for the Detection of Cancer versus Conventional Screening

To determine the relative sensitivity of the combination of two detection modalities, numbers of tumors that would have been found if only mammography and PE (conventional screening) had been performed were compared with numbers of tumors that would have been found if only mammography and US had been performed (Table 8). In women with dense breasts, the combination of mammography and US was significantly more sensitive (97.3%) than the combination of mammography and PE (74.7%; $P < .001$, Fisher exact test).

DISCUSSION

Mammography

Findings from our large prospective study of screening mammography, PE, and US show that sensitivity and specificity of screening mammography are 78% and 99%, respectively. For invasive cancer, the mammographic sensitivity is 73.0% for all breast types, but in the most dense breasts, it is only 44%. The mammographic sensitivity in women 49 years and younger was significantly lower (58%) than that in older women (83%). Others have reported a similar effect of age on mammographic sensitivity (34,35), but only recently has there been an examination of other factors possibly responsible for age-related sensitivity changes (22–24).

Our independent analysis of the effects of age, density, and hormonal status found that breast density was the most significant independent predictor of mam-

mographic sensitivity. In women with fatty breasts, the sensitivity of mammography was 98%; it decreased significantly to 48% in women with the highest density category breasts. This effect remained significant in women of all ages. There was no age-dependent difference in mammographic sensitivity in women with fatty breasts. However, in women with dense breasts, mammographic sensitivity in younger women was lower than that in older women. This independent age effect was much smaller than that of breast density.

Two theories have been advanced to explain why mammographic sensitivity is lower in younger women. First, younger women more often have dense breasts (30) in which it is harder to find smaller tumors in a background of dense fibroglandular tissue (17–20); as would be expected, women with interval cancers are more likely to have dense breasts (23). Second, tumors in younger women grow faster, resulting in more interval cancers (35–38). Our analysis of the mammographic appearance of cancer in different age groups failed to demonstrate a significant difference in tumor appearances to explain differences in sensitivity. Perhaps a larger sample size would have achieved this.

Our results differ from those of three published studies (21–23), in which independent effects of possible modifiers on mammographic sensitivity were analyzed. Kerlikowske et al (21) reported that in women younger than 50 years, breast density did not influence mammographic sensitivity. A possible explanation for their finding may be that findings from only nine patients of screened women younger than 50 years in whom cancers were missed at mammography were available for analysis, which reduced statistical power.

Rosenberg et al (22) found that in all age groups, women with dense breasts had a lower mammographic sensitivity than those with fatty breasts, but in women 50 years and older, this effect remained only when dense breasts were coupled with HRT. We found no independent effect of HRT on mammographic sensitivity. The decrease in sensitivity is accounted for by increased breast density alone, which is often, but not necessarily, linked to HRT. Rosenberg et al found that age was a minor determinant of mammographic sensitivity only in women 40 years and younger but that age made no difference in women 40 years and older. However, in their multicenter study, there was no systematic guideline for grading breast density; no report of how breast density was categorized; and incomplete reporting of density, estrogen use, and symptoms.

Finally Mandelson et al (23) reported in a large retrospective study of screened versus interval-detected cancers a consistent direct relationship between breast density and interval cancer risk, presumably those cancers that were detected in our study at US and/or PE. There was no independent age-related effect. Their finding may be due to the different methods used for ascertaining a false-negative result.

Interval Cancers

All previously reported study findings have relied on interval cancers (ie, cancers detected at some time after the index mammography) as the basis for defining false-negative mammograms. While calculation of interval cancer rates is important to optimize the interval between mammographic screenings and to identify false-negative examination findings, this method also has limitations. The time in which the interval cancer develops is completely arbitrary. The American College of Radiology has defined this interval as 12 months from the index mammography (24). In addition, the dependence on interval cancers makes the accuracy of cancer reporting crucial. If an index cancer is not reported to a cancer registry, false-negative examination findings will be missed, which will result in overestimation of mammographic sensitivity. We used the combination of multiple contemporaneous examinations to identify false-negative findings when tumors were identified with less than all the screening modalities. With this technique, it remains possible that in some patients, findings from all three examinations were simultaneously false-negative. An analysis of interval cancers in this population would have further diminished the sensitivities of each of the screening examinations.

US Screening

We report a large increase in detection of otherwise occult cancer if screening US is used in women with dense breasts. Forty-two percent of women with nonpalpable invasive cancer had their cancers detected only with screening US, and no other cancers were detected in these women with any other screening modality. Thirty-seven percent of all cancers in women with dense breasts were detected only with screening US. Seventy percent of cancers found only with screening US were subcentimeter, and 89% were node-negative, which conferred the best prognosis and the widest range of treatment options.

Interval cancers are at a more advanced stage at diagnosis and women have poorer stage-specific survival than women whose cancers are detected with mammography (39–45). Screening US seems to depict many of these cancers while they are at a smaller size and earlier stage than they would have been if they had been detected later as interval cancers. In our total population, 15% of cancers were detected only with US. This falls within the reported rates of interval cancers reported by others, which supports our theory that, if not detected with US, these cancers would grow to a larger size before becoming clinically apparent.

Tabar et al reported (33) that it is more difficult to detect tumors in dense breasts with screening mammography but concluded that the smaller mortality effect of screening in women aged 40–49 years is due to faster progression of some tumors in this age group and rapid increase in incidence during this decade. We did find an independent effect of age on mammographic sensitivity in women with similar density category breasts. The importance of finding these mammographically occult faster-growing tumors may have an effect on the overall survival. US alone enabled the diagnosis of cancer in 38% of all women younger than 50 years and with dense breasts and appears to be well suited for depicting these cancers before they become clinically apparent.

Some have argued that, much as with mammographic screening, nothing short of a randomized controlled trial with death as an endpoint is acceptable as proof of benefit from US screening (46). Our study findings show that cancers detected only with US are similar in size and stage to those detected only with mammography. Thus, it is likely that finding these early cancers at US would confer improved survival to the women in whom they are diagnosed. In addition to our previous study (30), there are three studies in which US was evaluated as a screening modality (47–49). The results of each of these studies closely approximate our findings.

The benefits of routine screening breast US must be weighed against the false-positive rate (2.4%) of this examination and the cost of the subsequent diagnostic procedures. This is comparable to, but

additional to, the negative biopsy findings prompted with screening mammography. However, the positive biopsy rate for masses detected only with US is lower than that for masses detected at mammography. This is in part, because US examination of mammographically and physically identified masses can help forestall biopsy of masses that have a benign appearance at US. To achieve the increased rate of cancer detection of small nonpalpable low-stage tumors with screening US, the overall false-positive rate in our study increased from 65.9% in women conventionally screened to 74.6% when screening US was added for women with dense breasts. However, since percutaneous US-guided biopsy is more cost-effective (50) and less invasive than and as effective as surgical biopsy, the benefit of a large increase in early stage cancer detection with screening US may outweigh the lower positive biopsy rate of screening US.

The prevalence of cancers detected with screening US is largest in the highest-risk women. If screening US were limited to high-risk women with dense breasts, nonpalpable invasive cancers would have been diagnosed in 26% more patients than were diagnosed with conventional screening versus 42% if the entire population of women with dense breast were screened; with an increase in the overall negative biopsy rate of only 3.6%.

PE Screening

There has never been a proven mortality benefit ascribed to the performance of breast palpation whether performed by physician or the patient.

In our study, only 28% of all cancers were detected at screening palpation, which had a specificity of 99.4%. Our reported sensitivity is lower than that reported by the United States National Screening Program (29), in which the sensitivity in asymptomatic women was reported as 36.1%, and is much lower than that reported by the Canadian National Breast Screening Study (CNBSS) (28), which reported sensitivities ranging from 57% to 83% for women aged 50–59 years and a sensitivity of 71% in women aged 40–49 years who underwent a single screening. However, in the former study, detection of interval cancers was only possible among women with more than one breast-screening record during the study interval, and only 26% of all participants were screened more than once. Therefore, the interval cancer rate was unusually low, which elevated the sensitivity of PE. In the CNBSS, the patient cohort did not exclude women who were symptomatic and included an unspecified number of women with self-reported lumps. This too, would elevate the sensitivity of screening clinical breast examination. Our data confirms that the sensitivity of PE is not influenced by age, and that sensitivity does not vary with breast density. We found that palpable cancers were significantly larger and higher stage than nonpalpable screening-detected cancers. Detection of cancers before they become clinically apparent is crucial.

In fatty and dense breasts, only 2% and 3%, respectively, of cancers were found at palpation that were not otherwise detected. Further, the combination of screening US and mammography would have resulted in significantly more cancers detected than the combination of screening PE and mammography. However, it is possible that PE of dense breasts, which enabled detection of lesions prior to the performance of breast US, may have been responsible for a more focused US examination of the palpable mass, which may have raised the apparent sensitivity of screening US. Therefore, PE with its small cancer yield, lack of a consistent ability to demonstrate smaller lower-stage cancers, and lack of a demonstrated or inferred survival or treatment-related benefit may not be effective to perform concurrently with mammography in women with fatty breasts. Women with dense breasts, especially those that are high risk, may benefit more from more frequent US examinations than from more frequent PEs.

Limitations

An important limitation of our study was examiner bias. A single examiner knew the results of each examination prior to performing the next one. This knowledge could guide the examiner to a lesion on the succeeding examination and falsely elevate the performance of the subsequent tests. In our study, mammography was followed by palpation and then, in women with dense breasts, US. Additionally, grading breast density is subjective; other examiners might grade breasts differently and screen a different population with US. However, even though examiner bias exists, our study is representative of how breast cancer is diagnosed; that is, results of preceding tests are generally known prior to performing the succeeding test, and this method can be considered the optimal way to conduct screening.

Second, our pilot study results led us to assume that no additional cancers would be found with screening US in fatty breasts; therefore, the false-negative rate for mammography was calculated with different means in women with fatty than in those with dense breasts. In women with fatty breasts, we calculated the mammographic false-negative rate based solely on the results of the concurrently performed PE, whereas in women with dense breasts the results of both PE and screening US were used.

Third, we calculated sensitivity on the basis of the use of multiple contemporaneous examinations to identify false-negative examination findings. Our method leads to the possibility that all three simultaneous examinations may have negative findings in patients with tumors. If tumors could be found in these patients with some other technique, the sensitivities we reported would diminish. However, results of 12-month surveillance for interval cancers as a method to detect false-negative examination findings indicate that mammographic sensitivity of 85% diminishes to 70% with lengthening of the surveillance interval to 24 months (23). Further, several randomized controlled trials of mammographic performance report the sensitivity as ranging from 68% to 88% (1–8). By using our technique, mammographic sensitivity is 78%, which suggests that the error rate of our technique is not significantly different from that established by using a different calculation method.

Finally, some analyses in patient subgroups were based on small sample sizes, which limited statistical power.

Our data support several conclusions. Breast density is the single most important predictor of mammographic sensitivity at any age. In fatty breasts, mammographic sensitivity is 98% and the performance of routine PE at the time of mammography has a very low yield. In women with dense breasts, the sensitivity of mammography diminishes to a low of 48% in women with the most dense breasts. If screening US is performed after mammography and PE in women with dense breasts, the yield of women diagnosed with nonpalpable invasive cancers is increased by 42%. US is significantly more sensitive than palpation in women of all ages with dense breasts. The substitution of screening US for PE, performed in conjunction with mammography, increases the sensitivity of cancer detection from 75% for those conven-

tionally screened to 97%. The additional cancers detected with screening US are similar in size and stage to those detected mammographically and are smaller and lower stage than those that are palpable. The rate of cancer detection with screening US is highest in those who are at high risk for developing breast cancer.

Acknowledgment: The authors thank Jane Fey, MPH, for her invaluable statistical assistance.

References

1. Shapiro S. Periodic screening for breast cancer: the health insurance plan project and its sequelae, 1963–1986. Baltimore, Md: Johns Hopkins University Press, 1988.
2. Andersson I, Aspegren K, Janzon L, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. BMJ 1988; 297:943–948.
3. Tabar L, Fagerberg C, Gad A, et al. Reduction in mortality from breast cancer after mass screening with mammography: randomized trial from the Breast Cancer Screening Working Group of the Swedish National Board of Health and Welfare. Lancet 1985; 1:829–832.
4. Roberts MM, Alexander FE, Anderson TJ, et al. Edinburgh trial of screening for breast cancer: mortality at seven years. Lancet 1990; 335:241–246.
5. Chamberlain J, Coleman D, Ellman R, et al. Sensitivity and specificity of screening in the UK trial of early detection of breast cancer. In: Miller AB, Jr, Chamberlain J, Day NE, et al, eds. Cancer screening. Cambridge, England: Cambridge University Press, 1991; 3–17.
6. Frisell J, Eklund G, Hellstrom L, et al. Randomized study of mammography screening: preliminary report on mortality in the Stockholm trial. Breast Cancer Res Treat 1991; 18: 49–56.
7. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study. I. Breast cancer detection and death rates among women aged 40 to 49 years. CMAJ 1992; 147:1459–1476.
8. Miller AB, Baines CJ, To T, et al. Canadian National Breast Screening Study. II. Breast cancer detection and death rates among women 50 to 59 years. CMAJ 1992; 147:1477–1488.
9. Peeters PHM, Verbeek ALM, Hendriks JHCL, van Bon MJH. Screening for breast cancer in Nijmegen: report of 6 screening rounds, 1975–1986. Int J Cancer 1989; 43:226–230.
10. Elwood JM, Cox B, Richardson AY. The effectiveness of breast cancer screening by mammography in younger women (serial online). Online J Curr Clin Trials 1993 Feb 25; Doc No 32.
11. Fletcher SW, Black W, Harris R, Rimer B, Shapiro S. Report of the international workshop on screening for breast cancer. J Natl Cancer Inst 1993; 85:1644–1656.
12. Tabar L, Fagerberg G, Chen H, et al. Efficacy of breast cancer screening by age: new results from the Swedish two-county trial. Cancer 1995; 75:2507–2517.
13. Tabar L, Fagerberg G, Duffy SW, Day NE. The Swedish Two-County Trial of mammographic screening for breast cancer: recent results and calculation of benefit. J Epidemiol Community Health 1989; 43:107–114.
14. Shapiro S. Periodic screening for breast cancer: the HIP Randomized Control Trial—Health Insurance Plan. J Natl Cancer Inst Monogr 1997; 22:27–30.
15. Alexander FE. The Edinburgh randomized trial of breast cancer screening. J Natl Cancer Inst Monogr 1997; 22:31–35.
16. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. Efficacy of screening mammography: a meta-analysis. JAMA 1995; 273: 149–154.
17. Holland R, Hendriks JH, Mravunac M. Mammographically occult breast cancer: a pathologic and radiologic study. Cancer 1983; 52: 1810–1819.
18. Bird RE, Wallace TW, Yankaskas BC. Analysis of cancers missed at screening mammography. Radiology 1992; 184:613–617.
19. Ma L, Fishell E, Wright B, Hanna W, Allan S, Boyd NF. Case control study of factors associated with failure to detect breast cancer by mammography. J Natl Cancer Inst 1992; 84: 781–785.
20. Feig SA, Shaber GS, Patchefsky A, et al. Analysis of clinically occult and mammographically occult breast tumors. AJR Am J Roentgenol 1977; 128:403–408.
21. Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Effect of age, breast density, and family history on the sensitivity of first screening mammography. JAMA 1996; 276: 33–38.
22. Rosenberg RD, Hunt WC, Williamson MR, et al. Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms in Albuquerque, New Mexico. Radiology 1998; 209:511–518.
23. Mandelson M, 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; 92:1081–1087.
24. American College of Radiology. Illustrated Breast Imaging Reporting and Data System (BI-RADS). 3rd ed. Reston, Va: American College of Radiology, 1998.
25. Feig SA, D'Orsi CJ, Hendrik RE, et al. American College of Radiology guidelines for breast cancer screening. AJR Am J Roentgenol 1998; 171:29–33.
26. Eastman P. NCI adopts new mammography screening guidelines for women. J Natl Cancer Inst 1997; 89:538–540.
27. Leitch AM, Dodd GD, Costanza M, et al. American Cancer Society guidelines for the early detection of breast cancer: update 1997. CA Cancer J Clin 1997; 47:150–153.
28. Baines CJ, Miller AB, Bassett AA. Physical examination: its role as a single screening modality in the Canadian National Breast Screening Study. Cancer 1989; 63:1816–1822.
29. Bobo JK, Lee NC, Thames SF. Findings from 752,081 clinical breast examinations reported to a national screening program from 1995 through 1998. J Natl Cancer Inst 2000; 92:971–976.
30. Kolb TM, Lichy J, Newhouse JH. Occult cancer in women with dense breasts: detection with screening US—diagnostic yield and tumor characteristics. Radiology 1998; 207:191–199.
31. 41st World Medical Assembly. Declaration of Helsinki: recommendations guiding physicians in biomedical research involving human subjects. Bull Pan Am Health Organ 1990; 24:606–609.
32. Beahrs OH et al, eds. Manual for staging of cancer. 3rd ed. Philadelphia, Pa: Lippincott, 1988.
33. Tabar L, Fagerberg G, Duffy SW, Day NE, Gas A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. Radiol Clin North Am 1992; 30:187–210.
34. Brekelmans CT, Collette HJ, Collette C, Fraschboud J, de Waard F. Breast cancer after a negative screen: follow-up of women participating in the DOM Screening Programme. Eur J Cancer 1992; 28A:893–895.
35. Gilliland FD, Joste N, Stauber PM, et al. Biologic characteristics of interval and screen-detected breast cancers. J Natl Cancer Inst 2000; 92:743–749.
36. Tabar L, Fagerberg C, Chen HH, Duffy SW, Gad A. Screening for breast cancer in women aged under 50: mode of detection, incidence, fatality and histology. J Med Screen 1995; 2:94–98.
37. Ikeda DM, Andersson I, Wattsgard C, Janzon L, Linell F. Interval carcinomas in the Malmo Mammographic Screening Trial: radiographic appearance and prognostic considerations. AJR Am J Roentgenol 1992; 159:287–294.
38. DeGroote R, Rush BF Jr, Milazzo J, Warden MJ, Rocko JM. Interval breast cancer: a more aggressive subset of breast neoplasias. Surgery 1983; 94:543–547.
39. Scroen AA, Wobbes T, van der Sluis RF. Interval carcinomas of the breast: a group with intermediate outcome. J Surg Oncol 1996; 63:141–144.
40. Senie RT, Lesser M, Kinne DW, Rosen PP. Method of tumor detection influences disease free survival of women with breast carcinoma. Cancer 1994; 73:1666–1672.
41. Frisell J, von Rosen A, Wiege M, Nilsson B, Goldman S. Interval cancer and survival in a randomized breast cancer screening trial in Stockholm. Breast Cancer Res Treat 1992; 24: 11–16.
42. Holmberg LH, Tabar L, Adami HO, Bergstrom R. Survival in breast cancer diagnosed between mammographic screening examinations. Lancet 1986; 2:27–30.
43. Vitak B. Invasive interval cancers in the Ostergotland Mammographic Screening Programme: radiological analysis. Eur Radiol 1998; 8:639–646.
44. Berkelmans CT, Peeters DeGroote R, Rush BF Jr, Milazzo J, Warden MJ, Rocko JM. Interval breast cancer: a more aggressive subset of breast neoplasias. Surgery 1983; 94:543–547.
45. Porter PL, El-Bastawissi AY, Mandelson MT, et al. Breast tumor characteristics as predictors of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 1999; 91:2020–2028.
46. Kopans DB. Breast-cancer screening with ultrasonography. Lancet 1999; 354:2096–2097.
47. Jacob D, Brombart JC, Muller C, Lefebvre C, Massa F, Depoerck A. Analyse des resultants de l'exerese chirurgicale de 137 lesions mammaires infracliniques: contribution de l'echographie dans le diagnostic precoce du cancer du sein. J Gynecol Obstet Biol Reprod 1997; 26:27–31.
48. Buchberger W, DeKoekkoek-Doll P, Springer P, Obrist P, Dunser M. Incidental findings on sonography of the breast: clinical significance and diagnostic workup. AJR Am J Roentgenol 1999; 173:921–927.
49. Kaplan SS. Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue. Radiology 2001; 221:641–649.
50. Liberman L, Feng TL, Dershaw DD, Morris EA, Abramson AF. US-guided core biopsy: use and cost-effectiveness. Radiology 1998; 208: 717–723.