Sydney Breast Imaging Accuracy Study: Comparative Sensitivity and Specificity of Mammography and Sonography in Young Women with Symptoms

Nehmat Houssami[1–3]
Les Irwig[2]
Judy M. Simpson[2]
Merran McKessar[1,4]
Steven Blome[1,5]
Jennie Noakes[4]

OBJECTIVE. We examined the age-specific sensitivity and specificity of mammography and sonography in symptomatic women to determine the age below which sonography may be the more accurate imaging test, which may guide the choice of initial breast imaging examination based on the woman's age.

MATERIALS AND METHODS. Four hundred eighty subjects were sampled from all women consecutively attending a symptomatic breast clinic between 1994 and 1996 and ranging in age from 25 to 55 years. We included all 240 women shown to have breast cancer (thus avoiding selection bias) and 240 age-matched women shown not to have cancer. Mammograms and sonograms were prospectively interpreted independently and without knowledge of age by two radiologists in a blinded manner, with a third radiologist arbitrating disagreements. Sensitivity and specificity of each imaging test in relation to age were examined using logistic regression modeling, and accuracy was compared using the chi-square test for paired proportions.

RESULTS. Sensitivity and specificity of each test were not linearly associated with age; however, the sensitivity of mammography increased substantially in women older than 50 years. Sonographic sensitivity of 81.7% was not significantly greater than mammographic sensitivity of 75.8% ($\chi^2_1 = 2.06$, $p = 0.15$). However, in women 45 years old or younger, the sensitivity of sonography was 13.2% (95% confidence interval, 2.1–24.3%) greater than that of mammography. The specificity of both tests was approximately 88.0%.

CONCLUSION. These data show that sonography is the more accurate imaging test in women 45 years old or younger who present with breast symptoms and may be an appropriate initial imaging examination.

Women who present with breast symptoms or who have palpable findings on clinical examination are usually investigated with breast imaging, which generally consists of mammography or breast sonography or both. The choice of primary breast imaging in examining women with symptoms is partly based on age. However, despite the importance of age in clinical practice, little evidence exists as to the appropriate age (or age range) that delineates the choice of initial diagnostic breast imaging in symptomatic women. In the absence of evidence, experts suggest that women younger than 35 years be examined with sonography, and women 35 years and older be examined with mammography, as the primary breast imaging modality [1].

A comprehensive review of the literature found little evidence about the comparative age-specific accuracy of mammography and breast sonography in symptomatic women [2]. A recent study has subsequently provided the first empiric evidence regarding the comparative sensitivity of both imaging tests in symptomatic women who underwent both examinations [3]. That study found that sonography was more sensitive than mammography in women younger than 62 years, the so-called crossover age, and mammography was more sensitive than sonography in women older than 62 years [3]. However, the study's authors acknowledged that the non-independent interpretation of the two tests and the analysis used may have led to underestimation of the sensitivity of mammography, and that the crossover age may be as early as 48 years [3]. For a valid comparison of the accuracy of two tests, the tests need to be interpreted independently (without knowledge of each other) in the same subjects [4].

Received May 13, 2002; accepted after revision September 3, 2002.

Supported by a health research grant from the Medical Benefits Fund of Australia.

[1]The Sydney-Square Breast Clinic, Medical Benefits Fund of Australia, Level 12, 97-99 Bathurst St., Sydney, New South Wales 2000, Australia.

[2]School of Public Health, Faculty of Medicine, A27, The University of Sydney, New South Wales 2006, Australia. Address correspondence to N. Houssami.

[3]NSW Breast Cancer Institute, Westmead Hospital, P. O. Box 143, Westmead, New South Wales 2145, Australia.

[4]Northern Sydney and Central Coast BreastScreen, The Cottage, Royal North Shore Hospital, St. Leonards, New South Wales 2065, Australia.

[5]Department of Diagnostic Radiology, Royal North Shore Hospital, St. Leonards, New South Wales 2065, Australia.

AJR 2003;180:935–940

0361–803X/03/1804–935

In this article, we report findings from the Sydney Breast Imaging Accuracy Study, the first study designed to establish the age-related accuracy of diagnostic mammography and sonography using independent (and blinded) interpretation of the two imaging tests. The study examines the comparative sensitivity and specificity of the two tests in symptomatic women 55 years old or younger. This study was of a clinically referred population and makes no inferences as to the accuracy or merits of screening mammography.

Materials and Methods

Subjects and Imaging

This retrospective study was reviewed and approved by the institutional ethics committee. Potential subjects included all women who were referred to a breast diagnostic center (clinic for symptomatic women) for the years 1994–1996 and who were 25–55 years old, inclusive. The selected age range of subjects was decided with consideration of the age commonly applied in clinical practice in deciding first-line imaging (30, 35, or 40 years). The maximum age of 55 years was chosen, in part, to ensure inclusion of most pre- and perimenopausal women, and because after age 55 mammography is the appropriate first-line imaging modality.

Cancer patients included all women in the defined inclusion criteria proven to have a breast malignancy. All women who were recommended to have surgery or who underwent surgery regardless of the clinic's recommendation, were followed up as standard monitoring procedure in the study center. Final histologic diagnosis was obtained for all patients who underwent surgical biopsy, and all cancer cases were verified by reviewing the histopathology report. Two cancer patients with typical malignant clinical findings and proven malignancy at cytology were not treated surgically but were not excluded from the study. Cancer cases included all types of invasive carcinoma as well as ductal carcinoma in situ. Eighteen eligible patients (7%) were excluded because images were not available.

To avoid verification bias [5, 6], subjects without cancer consisted of a randomly generated sample of consecutive women from the same population (irrespective of whether they had undergone biopsy) who were age-matched to cancer patients by year of birth. A subject was classified as not having breast cancer on the basis of the diagnostic assessment and on the basis of not being registered as having a breast cancer diagnosis at the New South Wales Central Cancer Registry during a 2-year interval after attending the study center. (Notification of malignant neoplasms is a statutory requirement in the state of New South Wales.) Two subjects without cancer were age-matched per each cancer patient. However, only one subject without cancer was included for each patient, using the following method to maximize the number of subjects without cancer from this randomly selected group who had undergone both tests. When both age-matched subjects without cancer had undergone only one test or both tests (27%), one was selected randomly by a coin toss. When only one of the two age-matched subjects without cancer had undergone both imaging tests (73%), that subject was chosen for inclusion. Because this method may theoretically lead to subjects without cancer having denser mammograms (and hence a higher specificity for mammography), we explored any potential effect this method may have in the analysis. The method has no potential influence on test sensitivity, however, because it applied only to subjects without cancer.

Medical records of all subjects were reviewed to verify eligibility, and all images performed at the time of diagnostic assessment were retrieved. In the study center, mammography is performed before sonography in all women older than 30 years who are referred for evaluation of symptoms. Sonography is performed in all women younger than 30 years and whenever a localized or focal finding (such as asymmetric thickening, lump) is noted on breast clinical examination regardless of the level of clinical suspicion. Mammography and sonography are almost always used in combination whenever some suspicion of malignancy exists regarding any individual component of the assessment. Mammography consisted of baseline mediolateral oblique and craniocaudal projections and all additional workup images obtained at the time of evaluation (≈ 60% of subjects had workup films). Workup films included films taken with a lead skin marker indicating the clinical area or modified projections that accommodate the clinical area. All sonographic images taken at the time of diagnosis were included. Sonography was performed with knowledge of the mammographic results by accredited sonographers with experience in breast sonography. It is routine practice in the study center to scan and include an image (in at least two planes) of the symptomatic area or the area palpated on clinical examination (as indicated by the examining clinician). The limitation of interpreting sonographic images on hard copy, rather than in real-time, is acknowledged and is explored further in the Discussion.

For the years included in the study, mammography was performed using one of three dedicated machines (Senograph 600T FD HF and 600T HF; General Electric Medical Systems, Milwaukee, WI), all of which used film-screen techniques. Sonographic equipment early in the study also included Aloka sonographic units (SSD-630; Aloka, Tokyo, Japan) with a 7.5-MHz transducer (mechanical sector probe). Two other units were used in 1995–1996: an Apogee 800 scanner (11.5-MHz linear array transducer; Advanced Technology Laboratories, Bothell, WA) and an SSA 250 scanner (7.5-MHz annular array transducer; Toshiba Medical Systems Division, Tokyo, Japan).

Prospective Reporting of Images in a Blinded Manner

For all subjects in the study, images (which were collected retrospectively) were prospectively interpreted by two radiologists who had no knowledge of the initial imaging reports or of the subjects' final diagnoses. The radiologists were aware of the study's aim and design but were unaware of the number and distribution of subjects and their ages. Disagreement between the two radiologists was resolved using a third radiologist, who was asked to arbitrate in disagreements [7] and given the other radiologists' opinions. All three radiologists had extensive experience in mammography and breast sonography and interpretation. Radiologists' reports were recorded on standardized forms using a categoric scale of 1–5 with an increasing level of suspicion (1 = no significant abnormality, 2 = benign, 3 = indeterminate, 4 = suspicious, and 5 = malignant). Radiologists were requested to interpret sonograms with consideration of published criteria [8].

To ensure independent interpretation of the two imaging tests and blinding of the radiologists performing the prospective interpretations, we used the following methods: the interpretation of each imaging test was maintained independent of knowledge of the other imaging test findings by completely separating mammographic and sonographic films, with each set of films being interpreted on separate days; films were presented to radiologists in batches of either mammography or sonography, with batches being interpreted at least 2–4 weeks apart; subjects' names and ages were deleted from images, thus eliminating any possibility of associating the mammographic films with the corresponding sonographic films; reviewers were separated (interpreting of films was performed by the radiologists on separate days to avoid any potential for discussion). Similarly, arbitration of disagreements was done on a separate day. Film interpretation throughout the phases of the study was performed in a separate room dedicated for this purpose, with restricted access while participating radiologists were interpreting.

Data Analysis

The prevalence of breast cancer in the study population was calculated. The overall sensitivity and specificity of each test were calculated. Categories 1 and 2 were considered negative for cancer, and categories 3–5, positive for cancer (in line with clinical practice in the study center, where imaging findings reported as categories 3–5 are recommended for further investigation). The accuracy of each test was also calculated in 5-year age groups and presented in tables and graphs. Sensitivity and specificity of the two tests were compared in all subjects and for the different age groups using McNemar's chi-square test for paired proportions (accuracy of the two imaging tests in subjects); the 95% confidence interval (CI) for the difference between paired proportions was also calculated.

Logistic regression analysis and cross-tabulations were performed using statistical software (version 6.12; SAS Institute, Cary, NC). The logarithm of the odds of a positive result was regressed on age (as a continuous variable) for cancer patients for each imaging modality to assess the influence of age on test sensitivity. The logarithm odds of a negative result were regressed on age for subjects without cancer for each imaging modality to assess the influence of age on test specificity. We cross-tabulated mammographic findings by sonographic findings in patients with cancer and subjects without cancer.

A post hoc analysis was performed to compare the sensitivity of the two tests in two age groups, older and younger than 45 years, using McNemar's chi-square test. We also tested whether the difference in sensitivity of the two tests was significant between the two age groups by testing the interaction between the test and the dichotomized age groups using logistic regression, taking into account the paired data using a generalized estimating equations model [9].

Results

The breast cancer prevalence in the study population was 2% (258 cancers from 13,583 women attending the symptomatic clinic). The study included 240 subjects with cancer (18 eligible patients were excluded because images were not available) who had all undergone both mammography and sonography, and 240 subjects without cancer who had undergone mammography, with 233 of those having undergone sonography. The histologic types of cancer in the 240 patients were invasive ductal (70%), ductal in situ (14%), invasive lobular (9%), tubular (4%), medullary (1%), and other types (1.2%), and no histology (0.8%). The frequency of symptoms (or reason for attending the symptomatic clinic) in all subjects is shown in Table 1. The mean age of subjects was 44.9 years, ranging from 27 to 55 years. Table 2 shows the comparative sensitivity of sonography and mammography in all subjects and in the different age groups. Although none of the age-specific differences for sensitivity reached statistical significance, the lower CI boundary is close to zero in both the 36–40 and 41–45 age groups. Table 3 shows the comparative specificity of the two tests in all subjects and in the different age groups. The sensitivity of the two tests in the different age groups is presented graphically in Figure 1, and the same data are presented for specificity in Figure 2.

Figure 1 shows that for women younger than 50 years, the sensitivity of both tests in relation to age has little variability. However, the sensitivity of mammography increases substantially after age 50. For women 46–55 years old, the sensitivity of the two tests becomes similar (Fig. 1) and the specificity of sonography and mammography are also similar (Fig. 2). Cross-tabulation of mammographic and sonographic results is shown in Tables 4 and 5. The sonographic sensitivity of 81.7% was 5.9% (95% CI, –1.5% to 13.2%) greater than the mammographic sensitivity of 75.8%. The difference in the sensitivities of the two imaging tests, however, was not statistically significant ($\chi^2_1 = 2.06$, $p = 0.15$).

For the 233 women without cancer who had both tests (Table 5), the specificity of

TABLE 1 Frequency of Presenting Symptoms or Reason for Attending Symptomatic Breast Clinic, All Subjects

Symptom	Cancer		No Cancer	
	No.	%	No.	%
Lump	128	53.3	76	31.7
Lump with discomfort	29	12.1	38	15.8
Thickening (± discomfort)	20	8.3	20	8.3
Localized lumpiness or nodularity	15	6.3	29	12.1
Second opinion (clinical or imaging)	14	5.8	4	1.7
No specific symptom[a]	12	5.0	38	15.8
Pain or localized discomfort	9	3.8	26	10.8
Nipple discharge (bloody)	3	1.3	0	0
Nipple or skin change (no discharge)	3	1.3	0	0
Lump with nipple or skin symptom	3	1.3	6	2.5
Nipple discharge (nonbloody)	2	0.8	3	1.3
Other	2	0.8	0	0
Total	240	100	240	100

[a]Women referred to diagnostic (symptomatic) clinic for perceived increased risk, history of breast disease, or investigation of symptom that had subsided by time of attendance.

TABLE 2 Comparative Sensitivity of Sonography and Mammography

Age Group (yr)	No. of Subjects	Sensitivity (%)		
		Sonography	Mammography	Difference (95% CI)
≤ 35	25	84.0	76.0	8.0 (–14.0 to 30.0)
36–40	39	84.6	69.2	15.4 (–2.8 to 33.6)
41–45	42	85.7	71.4	14.3 (–3.9 to 32.5)
46–50	83	77.1	75.9	1.2 (–11.5 to 13.9)
51–55	51	82.4	84.3	–1.9 (–16.8 to 12.9)
Total	240	81.7	75.8	5.9 (–1.5 to 13.2)

Note.—CI = confidence interval.

TABLE 3 Comparative Specificity of Sonography and Mammography

Age Group (yr)	No. of Subjects	Specificity (%)		
		Sonography	Mammography	Difference (95% CI)
≤ 35	27	85.2	92.6	–7.4 (–25 to 10.2)
36–40	37	91.9	83.8	8.1 (–0.7 to 16.9)
41–45	39	89.7	89.7	0 (–14.2 to 14.2)
46–50	83	86.8	85.5	1.3 (–7.9 to 10.4)
51–55	47	87.2	89.0	–2.2 (–14.6 to 10.4)
Total	233	88.0	87.6	0.4 (–5.0 to 5.8)

Note.—Specificity of mammography is based on 233 subjects without cancer who had both tests; specificity of mammography in all 240 subjects without cancer is 87.1%. CI = confidence interval.

sonography was 88.0%, which was not significantly different from the specificity of mammography (87.6%) ($\chi^2_1 = 0$, $p = 1$), with a difference in test specificity of 0.4% (95% CI, –5.0% to 5.8%). The specificity of mammography in women chosen for having had both tests was only 2% higher than the specificity in women randomly included, indicating that the method used for inclusion of subjects without cancer had little effect on estimates of specificity for mammography.

Table 4 shows that, of all cancers, 20% were correctly identified as cancer on sonography but not correctly identified as cancer on mammography, and 14% were correctly identified on mammography but not on sonography. Combining the two imaging tests, in which case imaging is considered positive if either test is positive, gives a high sensitivity of approximately 96%. That is, about 4% of cancers will be misdiagnosed when combining mammography and sonography in the investigation of symptomatic women. Table 5 shows that, of all women who did not have breast cancer, about 9% had false-positive findings on one test but were correctly identified as not having cancer on the other test. Combining the two tests, if imaging is considered negative when both tests have negative findings, gives a specificity of 79%. If both tests are combined in the investigation of symptomatic women, approximately 3% of women who do not have breast cancer will have a false-positive result on both mammography and sonography.

In women 45 years old or younger, the sensitivity of sonography was 84.9%, which was 13.2% (95% CI, 2.1%–24.3%) greater than the sensitivity of mammography (71.7%), indicating a significant difference in sensitivity for the two tests ($\chi^2_1 = 4.45$, $p = 0.035$). In women 46–55 years old, the sensitivity of sonography was 79.1%, which was the same as the sensitivity of mammography of 79.1% ($\chi^2_1 = 0$, $p = 1$). Testing for interaction between test type and dichotomized age group showed weak evidence ($p = 0.08$) of a difference between the two age groups in the difference in sensitivity of the two tests.

Regression of the probability of positive mammographic results on age in cancer patients (sensitivity) was not significant (likelihood ratio $\chi^2_1 = 1.23$, $p = 0.27$). Regression of the probability of a positive sonographic result on age in cancers was also not significant (likelihood ratio $\chi^2_1 = 0.16$, $p = 0.69$). For every 10 years' increase in age, the odds ratio for the probability of a positive mammographic result is 1.28 (95% CI, 0.83–1.97), and the odds ratio for the probability of a positive sonographic result is 0.91 (95% CI, 0.55–1.48). These results indicate that no significant linear relationship exists between age and the logarithm odds of a positive mammographic or sonographic result. Similarly, regression of the probability of a negative mammographic result on age in subjects without cancer (specificity) was not significant for mammography (likelihood ratio $\chi^2_1 = 0.06$, $p = 0.80$), and the odds ratio for a negative result for every 10 years' increase in age is 0.93 (95% CI, 0.53–1.64). Regression of the probability of a negative sonographic result on age in subjects without cancer was also not significant (likelihood ratio $\chi^2_1 = 0.19$, $p = 0.66$); and for every 10 years' increase in age, the odds ratio for a negative result is 0.88 (95% CI, 0.48–1.42).

Discussion

We used a comparative study design of higher quality than that of other studies published to date on the age-specific accuracy of the two imaging tests in symptomatic women (to our knowledge), which allowed calculation of the accuracy of each test independently of the other. Future studies can only improve on the methodology we have used by prospectively performing the imaging tests, thus ensuring that sonography is performed independently of any information from mammography, provided true independent interpretation is also maintained. We have minimized selection bias by including almost all cancers from consecutively attending women. We have avoided verification bias by using a random sample of age-matched women from the same population shown not to have breast cancer (and ensuring accurate classification of outcome).

The study contributes several findings to our knowledge of the accuracy of mammography and sonography in relation to age in women with breast symptoms. The greater sensitivity of mammography in women 50 years or older relative to younger women has been shown in other studies that have considered mammography only [10–12]. However, unlike some of these studies that show a progressive improvement in sensitivity with age, the sensitivity of mammography in our data is not linearly associated with age. The sensitivity of mammography in the Sydney Breast Imaging Accuracy Study remains relatively constant in women 45 years and younger and increases in women older than 45 years and, to a greater extent, in women 51–55 years.

Overall, the difference in the sensitivity of the two tests in all subjects is not statistically significant; however, in women 45 years or younger, sonography has a significantly greater sensitivity than mammography and a similar specificity. Although 45 years was a post hoc choice of age dichotomy, on the ba-

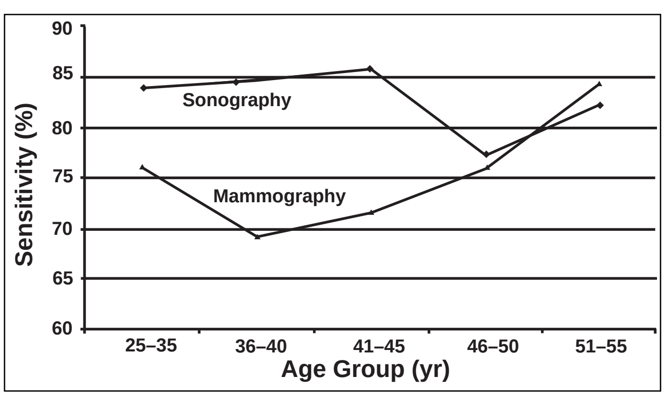

Fig. 1.—Graph shows sensitivity of sonography and mammography in different age groups.

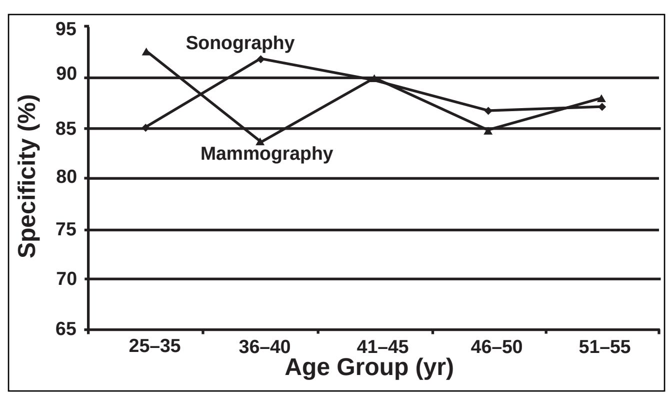

Fig. 2.—Graph shows specificity of sonography and mammography in different age groups.

sis of these data sonography may be the more appropriate initial test for investigating women with symptoms who are 45 years old or younger. It may be argued that the age of 45 years, as a means of deciding first-line breast imaging in women with symptoms, is later than expected. However, from a biologic perspective, this age as the decision basis makes sense because it correlates with the transition of hormonal (menopausal) status. In addition, despite differences in study design and analysis, this finding is in keeping with the recent study based on data from Florence that has also shown this crossover age to occur in the range of 48–62 years [3] and later than expected on the basis of expert opinion.

Our study also shows that there is little difference in the specificity of the two imaging tests, and that specificity is not influenced by age for either mammography or sonography. This fact may explain the different findings in published studies, with some reporting a greater specificity for sonography than for mammography [8, 11, 13], and others reporting a greater specificity for mammography than for sonography [14, 15]. Although seven of our subjects without cancer had not undergone sonography (< 2% of subjects), that fact is not likely to have any effect on comparative specificity because the proportion is very small and because specificity is approximately the same for both tests in these data. That fact also has no effect on test sensitivity because it does not apply to cancer patients and therefore does not invalidate our recommendations based on the differences in test sensitivity.

TABLE 4 Sensitivity of Mammography and Sonography in 240 Patients with Cancer

Mammography	Sonography		
	Positive (%)	Negative (%)	Total (%)
Positive (%)	148 (61.7)	34 (14.2)	182 (75.8)[a]
Negative (%)	48 (20.0)	10 (4.2)	58 (24.2)
Total (%)	196 (81.7)[a]	44 (18.3)	240 (100)

Note.—Percentages indicate proportion of all cases.
[a] Indicates sensitivity for each test.

TABLE 5 Specificity of Mammography and Sonography in 233 Subjects Without Cancer

Mammography	Sonography		
	Positive (%)	Negative (%)	Total (%)
Positive (%)	8 (3.4)	21 (9.0)	29 (12.5)
Negative (%)	20 (8.6)	184 (79.0)	204 (87.6)[a]
Total (%)	28 (12.0)	205 (88.0)[a]	233 (100)

Note.—Percentages indicate proportion of all cases.
[a] Indicates specificity for each test.

It may be argued that sonography is a real-time imaging test and any judgment of its accuracy should be based on real-time or dynamic interpretation, and that using hard-copy sonographic images in this study may lead to underestimation of the accuracy of sonography. Although this potential effect is acknowledged, dynamic features basically consist of compressibility and, to a lesser extent, mobility, which are not considered primary features of breast lesions seen on sonography [16]. We also point out that research undertaken to define breast sonographic criteria has similarly used sonographic hard-copy images [17]. This potential effect may be balanced by other factors inherent in sonography that may potentially overestimate its accuracy. Because sonography was performed with knowledge of the clinical opinion and the mammographic result, the operator may have been influenced by the level of suspicion or other information obtained from these tests. The hard-copy image is also likely to be the most representative image of the lesion obtained by systematic scanning. In summary, we believe that using a hard copy of sonographic films gives a realistic and valid estimate of test accuracy. Although using hard-copy sonograms prevents assessment of the dynamic features of lesions, which may potentially underestimate accuracy, this effect is likely to be balanced by a potential to slightly overestimate the performance of sonography by presenting the most representative image of the lesion.

Our study shows that combining both mammography and sonography has a greater sensitivity (96%) than either sonography alone (81.7%) or mammography alone (75.8%), although combining reduces specificity. These findings are in keeping with other studies of the accuracy of combined mammography and sonography, which have been extensively discussed in a recent paper that included an appraisal of published work on this topic [18].

Using the age of 45 years to decide the choice of first-line imaging in women with breast symptoms raises the question of whether mammography is appropriate as second-line imaging in women younger than 45 years. Apart from the individual clinical situation that guides decision making, the additional use of mammography may be decided on the basis of the sonographic result and the prevalence of breast cancer in the population being investigated (which is an indicator of the woman's probability of having the disease before test findings). If the sonographic result is positive, most clinicians would agree that mammography is indicated to define the extent of disease, in addition to contributing further diagnostic information. However, if the sonographic findings are negative, the benefit of mammography is related to the remaining breast cancer prevalence, once sonography has identified (and therefore removed) most cancers.

As an example, in a population of 1000 women younger than 45 years and with symptoms, in which the initial breast cancer prevalence is 2%, sonography will identify approximately 85% of cancers, so the remaining breast cancer prevalence is 0.3%. Mammography in this context would identify two additional cancers not identified on sonography and cause an additional false-positive result in 85 women who do not have breast cancer. In a population of 1000 women with symptoms in whom the initial breast cancer prevalence is 10%, mammography would identify 11 additional cancers not identified on sonography and cause an additional false-positive result in 78 women who do not have breast cancer.

We have shown that breast sonography is more accurate than mammography in symptomatic women 45 years or younger and may be an appropriate initial imaging test in those women. We recognize that the evidence from this research is contrary to current thinking and clinical practice in breast imaging of symptomatic women, and we recommend further research that examines this issue using a prospective study. Future research should also examine how hormonal status (rather than age alone) influences the accuracy of both tests. We suggest that any change from current practice should be introduced initially in the context of dedicated breast services, in which systematic sonographic scanning and monitoring of accuracy are routine practice, before wider implementation.

References

1. Dixon JM, Mansel RE. Symptoms, assessment, and guidelines for referral. In: Dixon JM, ed. *ABC of breast diseases*, 2nd ed. London: BMJ, **2000**:3–7
2. Irwig L, Macaskill P. *Evidence relevant to guidelines for the investigation of breast symptoms.* Sydney, Australia: National Breast Cancer Centre, **1997**:5–18
3. Houssami N, Ciatto S, Irwig L, Simpson JM, Macaskill P. The comparative sensitivity of mammography and ultrasound in women with breast symptoms: an age-specific analysis. *Breast* **2002**;11:125–130
4. National Health and Medical Research Council. *How to review the evidence: systematic identification and review of the scientific literature*. Canberra, A. C. T., Australia: Commonwealth of Australia, **2000**:62–63
5. Jaeschke R, Guyatt G, Sackett DL. Users' guide to the medical literature: how to use an article about a diagnostic test—are the results of the study valid? *JAMA* **1994**;271:389–391
6. Houssami N, Irwig L. Likelihood ratios for clinical examination, mammography, ultrasound and fine needle biopsy in women with breast problems. *Breast* **1998**;7:85–89
7. Mucci B, Athey G, Scarisbrick G. Double read screening mammograms: the use of a third reader to arbitrate on disagreements. *Breast* **1999**;8:37–39
8. Stavros AT, Thickman D, Rapp CL, et al. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. *Radiology* **1995**;196:123–134
9. Zeger SL, Liang K-Y, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics* **1988**;44:1049–1060
10. Sibbering DM, Burrell HC, Evans AJ, et al. Mammographic sensitivity in women under 50 years presenting symptomatically with breast cancer. *Breast* **1995**;4:127–129
11. Ciatto S, Rosselli del Turco M, Catarzi S, Morrone D. The contribution of ultrasonography to the differential diagnosis of breast cancer. *Neoplasma* **1994**;41:341–345
12. Dixon JM, Anderson TJ, Lamb J, Nixon SJ, Forrest APM. Fine needle aspiration cytology, in relationships to clinical examination and mammography in the diagnosis of a solid breast mass. *Br J Surg* **1984**;71:593–596
13. Negri S, Bonetti F, Capitanio A, Bonzanini M. Preoperative diagnostic accuracy of fine-needle aspiration in the management of breast lesions: comparison of specificity and sensitivity with clinical examination, mammography, echography, and thermography in 249 patients. *Diagn Cytopathol* **1994**;11:4–8
14. Moss HA, Britton PD, Flower CDR, Freeman AH, Lomas DJ, Warren RML. How reliable is modern breast imaging in differentiating benign from malignant breast lesions in the symptomatic population? *Clin Radiol* **1999**;54:676–682
15. Smallwood J, Guyer P, Dewbury K, et al. The accuracy of ultrasound in the diagnosis of breast disease. *Ann R Coll Surg Engl* **1986**;68:19–22
16. Tohno E, Cosgrove DO, Sloane JP. *Ultrasound diagnosis of breast diseases.* Edinburgh: Churchill Livingstone, **1994**:50–72
17. Rahbar G, Sie AC, Hansen GC, et al. Benign versus malignant solid breast masses: US differentiation. *Radiology* **1999**;213:889–894
18. Houssami N, Irwig L, Loy C. Accuracy of combined breast imaging in young women. *Breast* **2002**;11:36–40