



Interpretation of automated breast ultrasound (ABUS) with and without knowledge of mammography: a reader performance study

Per Skaane¹, Randi Gullien², Ellen B. Eben², Merete Sandhaug³,
Ruediger Schulz-Wendtland⁴ and Frank Stoeblen⁵

Abstract

Background: Automated breast ultrasonography (ABUS) has the potential to be an important adjunct to mammography in women with dense breasts.

Purpose: To compare reader performance and inter-observer variation of ABUS alone and in combination with mammography.

Material and Methods: This retrospective study had ethical committee approval. All women gave written informed consent. One hundred and fourteen breasts in 90 women examined by digital mammography and ABUS were interpreted by five radiologists using BI-RADS categories. The 114 breasts included 38 cancers and 76 normal or benign findings. In the first reading session ABUS only was interpreted, and in the second ABUS plus digital mammography. Image interpretations were done without knowledge of clinical or imaging results. A consensus panel analyzed false negative and false positive interpretations. Reading time was recorded for one radiologist. AUC was used for performance measurement, and kappa statistic for inter-observer variability.

Results: Mean size for cancers was 16.2 mm; area under the curve (AUC) values for ABUS alone and for combined reading were, respectively: reader A, 0.592–0.744; reader B, 0.740–0.947; reader C, 0.759–0.823; reader D, 0.670–0.688; reader E, 0.904–0.923; and all readers combined 0.730–0.823. The higher AUC for combined reading was statistically significant ($P < 0.05$) for reader B and for all readers. There was a considerable inter-observer variability. Observer agreement revealed following kappa values for ABUS alone and combined reading, respectively: reader A, 0.22–0.30; reader B, 0.33–0.44; reader C, 0.32–0.39; reader D, 0.07–0.14; and reader E, 0.34–0.43. Shadowing from dense parenchyma was the most common cause of false positive ABUS interpretations. Mean interpretation time for a bilateral normal ABUS examination was 9 min.

Conclusion: Observer agreement was higher and all radiologists improved diagnostic performance using combined ABUS and mammography interpretation. Combined reading should be standard if ABUS is implemented in screening of women with dense breasts.

Keywords

Automated breast ultrasonography, hand-held ultrasound, breast cancer screening, mammography, inter-observer variability

Date received: 14 August 2013; accepted: 1 March 2014

Introduction

Mammography screening has proven effective in reducing mortality due to early diagnosis of small breast cancers. Randomized controlled trials have demonstrated a significant decrease in breast cancer mortality

¹Department of Radiology, Oslo University Hospital Ullevaal, University of Oslo, Norway

²Department of Radiology, Oslo University Hospital Ullevaal, Oslo, Norway

³Department of Radiology, Østfold Sentralsykehus, Fredrikstad, Norway

⁴Department of Radiology, University of Erlangen, Erlangen, Germany

⁵Department of Radiology, Kliniken-Essen-Mitte, Essen, Germany

Corresponding author:

Per Skaane, Department of Radiology, Breast Imaging Center, Oslo University Hospital Ullevaal, Kirkeveien 166, N-0407 Oslo, Norway.
Email: PERSKA@ous-hf.no

(1). A recent study demonstrated 43% reduction among women attending a national screening program (2). Mammography is the only imaging test proven to reduce breast cancer mortality.

Sensitivity of mammography depends on breast density. Studies on women with dense breasts have demonstrated sensitivity less than 50% (3). More than 50% of the women younger than 50 years and at least one-third aged more than 50 years have breast density BI-RADS 3 and 4 (4). Several imaging tests have been recommended for women with dense breasts, of which two modalities have been implemented for screening: hand-held ultrasonography (HHUS) and magnetic resonance imaging (MRI).

HHUS is widely available and well tolerated. Screening with HHUS has found significantly more early-stage breast cancers (4–10). A limitation is that HHUS is time-consuming. MRI has a high sensitivity (3). Limitations for implementing MRI are high costs and availability. Costs and examination time can be reduced using “short first-pass MRI” technique (11). Neither HHUS nor MRI are realistic alternatives as adjuncts to mammography in population-based screening.

Automated breast ultrasound (ABUS) was introduced many years ago, but image quality obtained with old generation whole-breast scanners was inferior as compared with high-resolution HHUS. Modern scanners using high frequency transducers have overcome this limitation. Three-dimensional (3D) data sampling of ABUS makes it possible to obtain transverse, sagittal, as well as coronal images. Standardized images make ABUS reproducible and suitable for follow-up. ABUS is carried out by radiographers, and interpretation at dedicated workstation makes independent double reading in batch mode possible. Thus, ABUS possess the essential prerequisites for implementation in high-volume screening.

The purpose of our study was to evaluate diagnostic performance and inter-observer variability using ABUS alone *versus* combined ABUS and mammography interpretation.

Material and Methods

Patients

Institutional review board approved the study, and all patients provided informed written consent. Women referred to the Breast Imaging Center were invited to participate when daily work-load permitted additional ABUS. All women underwent scheduled imaging first, and afterwards ABUS examination.

Ninety women were included. Indication for imaging was a palpable mass in 31 (35%), work-up of suspicious

findings at mammography in 47 (52%), surveillance in nine (10%), and nipple discharge in three (3%). Both breasts were examined with ABUS in 24 and one breast in 66 patients. For purpose of the study, one breast was defined as one case. Thus, a total of 114 breasts (one breast = one case) were included. Mean age of the 90 patients was 54.9 years (range, 20–86 years; median age, 57.5 years).

Final diagnosis was based on needle biopsy or surgical histology. Patients with normal or benign findings dismissed without biopsy had a 1-year follow-up. A total of 76 cases were normal or benign: 45 normal, 10 cysts, 15 fibroadenomas, two papillomas, one phyllodes tumor, and three radial scars. Histology of 38 malignant tumors included 25 invasive ductal carcinomas (IDC), six mixed IDC and ductal carcinoma in situ (DCIS), one pure DCIS, five invasive lobular carcinomas (ILC), and one invasive apocrine carcinoma. Mean size of the 38 cancers was 16.2 mm (range, 4–50 mm; median, 14.0 mm); mean size of the 24 screening-detected cancers was 13.2 mm (range, 4–25 mm; median, 11.0 mm); and mean size of the 14 cancers presenting as palpable lump was 21.4 mm (range, 8–50 mm; median, 18.5 mm). Mean age of cancer patients was 61.3 years (range, 41–86 years) and median age was 62.5 years.

Parenchymal density was grouped according to Breast Imaging Reporting and Data System (BI-RADS) as category 1 (fatty) or 2 (scattered dense) in 56 cases, and as density 3 (heterogeneously dense) or 4 (extremely dense) in 48 cases. In 10 cases (seven normal and three fibroadenomas) mammography was not carried out since this had been done at another institution.

Final diagnoses and densities of the 114 cases, given as mean BI-RADS score by the five readers, are summarized in Table 1.

Mammographic and ultrasonographic examinations

Full-field digital mammography (FFDM) was carried out using a Senographe DS unit (GE Healthcare, Milwaukee, WI, USA). Examinations included two standard (cranio-caudal and medio-lateral oblique) views. Additional (latero-medial, spot compression, and magnification) views were performed if indicated but not offered for the study interpretation sessions.

HHUS was carried out using a high-resolution sonographic unit with a 12 MHz linear transducer (Aplio 80 System, Toshiba Medical Systems, Tokyo, Japan). All ultrasonographic examinations were performed by one of five in-house radiologists. HHUS images were not offered for the study interpretation sessions.

Automated breast volume scanning (ABUS) was carried out using a ABUS unit (SomoVuTM, U-System, Inc., San Jose, CA, USA). The scanner is

Table 1. Breast parenchyma density (ACR, BI-RADS categories 1–4) and final diagnoses for the 114 automated whole breast (volume) ultrasound examinations.

Density BI-RADS	Benign: normal	Benign: cyst	Benign: fibroadenoma	Benign: other tumors	Benign: radial scar	Cancer: IDC	Cancer: IDC + DCIS	Cancer: ILC/others	Total
1	3	0	0	1	1	1	2	0	8
2	23	5	3	0	0	14	1	2	48
3	11	3	5	2	1	9	2	4	37
4	1	2	4	0	1	1	2*	0	11
No mx	7	0	3	0	0	0	0	0	10
Total	45	10	15	3	3	25	7	6	114

*One case of pure DCIS.

DCIS, ductal carcinoma in situ; IDC, invasive ductal carcinoma; ILC, invasive lobular carcinoma; mx, mammography.

equipped with a linear 7.5–10 MHz transducer. Four radiographers trained with ABUS performed all examinations. Patients were examined in supine position with the ipsilateral arm above the head. Three views were usually performed for women with small and medium-size breasts: A “frontal” including areola area, a frontal-medial to cover medial quadrants, and a frontal-lateral view to include lateral quadrants. In women with large breasts, additional views were occasionally needed. The examination technique has been described earlier (12). Two views were performed in 42 (37%), three views in 51 (45%), four views in 20 (17%) breasts, and five views in one case.

Image study interpretations

Five radiologists with 5–29 years of experience in breast imaging participated in the study: two in-house radiologists (readers C and E with 25 and 29 years of experience in clinical mammography, respectively) and three external readers of which one (reader B, with 5 years of experience) was domestic and two were from abroad (readers A and D, with 22 and 15 years of experience, respectively). Two of the three domestic radiologists (reader B and C) had no experience with ABUS prior to the study, and 4 hours of training with ABUS interpretation and how to operate the dedicated ABUS review workstation was given before study began. The radiologists analyzed all images in both reading sessions without knowledge of patient history, clinical findings, results of initial imaging work-up, or histology reports. Images were anonymized and presented to the readers in the same order (sequence). The cases were randomly mixed by one of the authors (RG) who did not take part in interpretations in such a way that the two breasts from patients with bilateral examination always were separated by several other cases. The two in-house and the domestic external radiologists interpreted all ABUS cases at a dedicated review

workstation located in a darkened room where also the dedicated FFDM workstation was located. The anonymized mammography and ABUS examinations were sent to the two foreign radiologists using a mini hard-disk since both had dedicated FFDM and ABUS workstations available.

The first reading session 10 weeks after the last woman had been examined included ABUS only. The domestic readers were recommended to use the in-house reading protocol, which included three main steps. First, analyzing transverse images slice-by-slice and compare suspicious findings with the corresponding reconstructed sagittal views. If a suspicious finding was noticed, it was recommended to put a marker on it and wait with interpretation until the end (see below). Second, coronal images were analyzed looking for “holes” (“black” or “white” depending on setting) and the “retraction phenomenon”. Third, the final analysis of suspicious lesions, if any. This reading protocol was a recommendation, and its use was optional. The two foreign radiologists had both experience with ABUS, and no reading protocol was recommended.

The second reading session including combined ABUS and mammography was carried out 8–12 weeks after the first. The readers had to return their completely filled out sheets from the first session before the anonymized ABUS and FFDM were offered for combined interpretation. The readers were again blinded to clinical, imaging and pathology results, and to their ABUS interpretations from the first session. The author (RG) not taking part in interpretation again mixed images in such a way that cases from women having bilateral examination were always separated by several cases from other women. The sequence of cases was different from the first reading session, but similar for all readers.

In both sessions the radiologists had to fill out a sheet for each case giving their BI-RADS score: 1, normal; 2, benign; 3, probably benign; 4, suspicious;

and 5, highly suggestive of cancer. BI-RADS subgroups 4A–C was not used. In the first session, the readers should specify their main ultrasound diagnosis (normal; cyst; differentiation cyst-tumor not possible; benign tumor; indeterminate tumor; malignant tumor; other abnormality), and size (in mm) of a tumor. The readers were asked to answer “yes” or “no” whether the ABUS examination was properly done, and if “no” give the reason why inadequate (breast incompletely examined; artefacts, or specify if other reasons). The sheets were checked by the radiographer (RG) for completeness before the second reading session could start. In the second session, the radiologists had to give assessment score for combined interpretation using the same 5-point BI-RADS scale, and to specify mammographic findings (normal; mass; distortion; asymmetric density; calcifications; or a combination of findings). Only final BI-RADS assessment scores were used for comparing results in the two reading sessions.

In the second session, each reader was asked to give the mammographic density score according to the BI-RADS scale (1, fatty breast with glandular density less than 25%; 2, scattered fibroglandular density 25–50%; 3, heterogeneously dense breasts with 51–75% glandular tissue; and 4, extremely dense breast with more than 75% glandular tissue). Mean scores given by the five readers are shown in Table 1.

Interpretation time for the 114 ABUS examinations was recorded for one reader (PS). The organizing radiographer (RG) manually recorded reading time for each case using a stop-watch.

After finishing both sessions, the three domestic radiologists in a consensus meeting analyzed ABUS interpretations with false positive and false negative results. All sheets filled out by the radiologists in both sessions, clinical history, imaging findings, and pathology reports were available at this meeting.

Statistical analysis

Sensitivity and specificity were defined using dichotomized BI-RADS assessment scores: Scores 1–3 were considered normal/benign and grouped “negative”, whereas BI-RADS scores 4–5 were defined as suspicious and grouped as “positive”. Diagnostic performance for individual readers for both sessions was calculated using receiver operating characteristic (ROC) analysis, and area under the ROC curve (AUC) was used for comparison (EpiInfo Version 3.3).

Kappa statistic was used for comparison of interobserver variability using the 5-point BI-RADS assessment categories. Unweighted kappa values are presented for a three-level scale of the collapsed 5-point assessment scale. For this 3-level scale, BI-RADS categories 1 and 2 were grouped as “benign”,

category 3 remained “indeterminate”, and BI-RADS categories 4 and 5 were grouped as “suspicious/malignant”. Collapsing the BI-RADS 5-point assessment scale into a three-level rating scale seems reasonable since the differentiation between normal and definitely benign on one hand, and suspicious and malignant on the other hand is trivial and of no clinical consequence.

Results

ABUS only interpretation

Excluding non-neoplastic lesions (cysts and radial scars) from analysis, a total of 18 benign tumors were diagnosed. Detection rate for these 18 tumors was 39% (7/18) for reader A, 56% (10/18) for reader B, 61% (11/18) for reader C, 100% (18/18) for reader D, and 72% (13/18) for reader E.

Thirty-eight cancers were included (Table 1). In the consensus review three malignant lesions were considered not to be visible on ABUS: one Paget’s disease, one pure DCIS, and one mixed DCIS and IDC. Excluding these three non-visible malignant lesions from analysis, reader A diagnosed a suspicious or malignant tumor in 34% (12/35), reader B in 66% (23/35), reader C in 77% (27/35), reader D in 100% (35/35), and reader E in 91% (32/35) of cases. The detection of benign and malignant tumors on ABUS only for the five readers is presented in Fig. 1.

ROC analysis based on the 5-point BI-RADS categories for benign and malignant tumors showed that the area under the curve (AUC) varied from 0.592 to 0.904 among the radiologists (Table 2).

The false positive rate (BI-RADS score 4 or 5 for a benign lesion) was for reader A 4% (3/76), for reader B

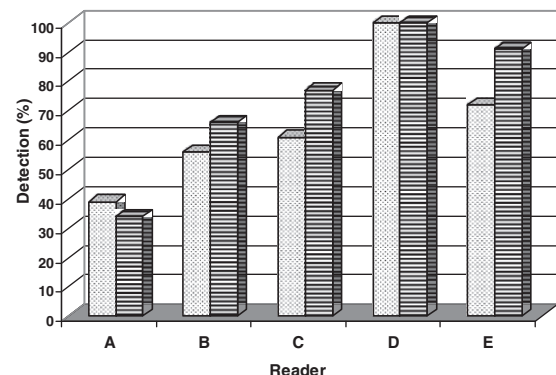


Fig. 1. Bar graph showing the percentage of detection (perception) rate for the 18 benign (dotted columns) and the 35 malignant (columns with horizontal lines) tumors by each of the five radiologists using ABUS only in the first reading session. Three malignant lesions not visible on ABUS are excluded from analysis.

11% (8/76), for reader C 14% (11/76), for reader D 71% (54/76), and for reader E 14% (11/76). In 25 cases two or more readers had a false positive score. Retrospective consensus evaluation of these 25 cases concluded that the probably reasons for false positives were a benign tumor (fibroadenoma) misinterpreted as suspicious in 60% (15/25) of cases, shadowing from heterogeneous parenchyma in 24% (6/25) of cases, shadowing from scarring tissue in 8% (2/25), and there was no explanation for the false positive score by two readers in 8% (2 cases).

False negative interpretation by two or more readers occurred in 18 malignant cases. At consensus

evaluation, shadowing was the explanation for the missed cancers in six (33%) of these cases, whereas the tumor was overlooked in nine (50%) cases. Three cancers were not visible at ABUS.

Observer agreement using unweighted kappa for the three-level scale of the collapsed 5-point BI-RADS categories revealed following mean kappa values with the four comparisons of each radiologists with the others (individual values in parenthesis): reader A, 0.22 (0.32, 0.26, 0.04, 0.27); reader B, 0.33 (0.32, 0.45, 0.05, 0.49); reader C, 0.32 (0.26, 0.45, 0.07, 0.50); reader D, 0.07 (0.04, 0.05, 0.07, 0.12); and reader E, 0.34 (0.27, 0.49, 0.50, 0.12). The kappa values are presented in Table 3.

Table 2. Diagnostic performance (ROC analysis) of the five radiologists and the overall pooled results based on the 5-point BI-RADS assessment scale. The AUC with 95% CI are presented for ABUS only and for combined ABUS and mammography interpretation.

Reader	ABUS only AUC (95% CI)	ABUS + Mammography AUC (95% CI)	Significance
A	0.592 (0.475–0.709)	0.744 (0.638–0.851)	n.s.
B	0.740 (0.629–0.852)	0.947 (0.900–0.994)	Significant
C	0.759 (0.652–0.867)	0.823 (0.728–0.919)	n.s.
D	0.670 (0.571–0.769)	0.688 (0.592–0.784)	n.s.
E	0.904 (0.828–0.980)	0.923 (0.857–0.988)	n.s.
All	0.730 (0.682–0.779)	0.823 (0.783–0.863)	Significant

Table 3. Kappa scores on agreement between the five radiologists for ABUS only interpretation (K3abus) and combined ABUS and mammography reading (K3 mx). The mean kappa score is the mean value of agreement in the four comparisons of each radiologist with the others.

Reader	Reader				Mean kappa
	B	C	D	E	
A	K3abus = 0.32 K3 mx = 0.46	K3abus = 0.26 K3 mx = 0.31	K3abus = 0.04 K3 mx = 0.07	K3abus = 0.27 K3 mx = 0.35	0.22 0.30
B		K3abus = 0.45 K3 mx = 0.51	K3abus = 0.05 K3 mx = 0.15	K3abus = 0.49 K3 mx = 0.64	0.33 0.44
C			K3abus = 0.07 K3 mx = 0.18	K3abus = 0.50 K3 mx = 0.56	0.32 0.39
D				K3abus = 0.12 K3 mx = 0.15	0.07 0.14
E				K3abus K3 mx	0.34 0.43

K3abus: Kappa value on the three-level scale of the collapsed 5-point rating scale for ABUS interpretation only. K3 mx: Kappa value on the three-level scale for combined ABUS and mammography reading.

For both reading modes, the 5-point BI-RADS assessment scale has been collapsed into a 3-point scale: Assessment categories 1 and 2 as one category, assessment category 3 unchanged, and assessment categories 4 and 5 merged into one category.

Combined FFDM and ABUS interpretation

The number of malignant tumors having a positive BI-RADS score 4 or 5 using combined reading as compared with ABUS alone increased for four of the five radiologists and was unchanged for one: reader A, 10 vs. 19; reader B, 24 vs. 35; reader C, 27 vs. 28; reader D, no change; and reader E, 32 vs. 33.

ROC analysis showed that AUC for combined interpretation increased as compared with ABUS alone for all five radiologists. The increase was statistically significant for one radiologist (reader B). Overall pooled AUC values also revealed a significant increase using combined reading as compared with ABUS alone (Table 2).

Observer agreement using the three-level unweighted kappa statistic of the collapsed 5-point BI-RADS assessment revealed following mean kappa values for each radiologist (the individual kappa values for comparison with the others in parenthesis): reader A, 0.30 (0.46, 0.31, 0.07, 0.35); reader B, 0.44 (0.46, 0.51, 0.15, 0.64); reader C, 0.39 (0.31, 0.51, 0.18, 0.56); reader D, 0.14 (0.07, 0.15, 0.18, 0.15); and reader E, 0.43 (0.35, 0.64, 0.56, 0.15). The kappa values are summarized in Table 3.

Interpretation time

Mean interpretation time for the 114 ABUS examinations (one examination = one breast) recorded manually for one of the readers (PS) was 276 s (i.e. 4 min 36 s) per breast, corresponding to about 9 min for a bilateral examination. The reading time varied depending on the presence of abnormalities. Thus, the mean interpretation time per breast was 244 s for 55 normal/benign breasts (normal and fibrocystic changes), 290 s for 21 breasts containing benign tumor, and 316 s (i.e. 5 min 16 s) for breasts with a malignant tumor detected by the reader.

Discussion

The main findings of this study are the considerable inter-observer variation for interpretation of ABUS, and the improvement of reader agreement and diagnostic performance for combined ABUS and mammography reading as compared with ABUS alone. These results are in accordance with studies comparing HHUS *versus* combined reading (13). Our results indicate that ABUS reading should be performed with knowledge of mammographic findings and by the same radiologist. This important issue has so far received little attention.

The incremental cancer detection rate using supplemental HHUS has been two to four cancers per 1000 women screened (4,6,7,10,14,15). Similar increase in

diagnostic yield was also reported using automated whole breast ultrasound (16). HHUS is time-consuming. To the best of our knowledge, only Austria has implemented ultrasound in organized screening. Data from Austria show that 38% of attending women (having density BI-RADS 3 and 4) underwent additional ultrasound (17).

Personalized (individualized) screening with supplemental ultrasound for women at average risk for breast cancer is a hot topic. The issue has been actualized after a law was enacted in the US state of Connecticut in October 2009, requiring radiologists to communicate breast density information to patients undergoing mammography, and mandating that women are informed of the possible benefit of ultrasound as additional screening modality (14,18). Meanwhile further US states have passed similar laws.

A problem with supplemental ultrasound is the high number of false positive findings. Reported positive predictive values for ultrasound-detected lesions have been in the range of 5.5–10.3% (4,10,14,18). Training might improve performance for ultrasound since high negative predictive values have been reported using systematic tumor descriptor analysis (5,19,20).

ABUS possesses prerequisites for implementation in population-based screening. ABUS is carried out by radiographers since breast density can be automatically determined at FFDM acquisition station (21). Consequently, supplemental ultrasound can be performed during a single visit. Standardized images can be interpreted on dedicated viewing stations using independent double reading in batch mode. ABUS has shown promising results (16), and most cancers larger than 10 mm are demonstrated in transverse and reconstructed sagittal and coronal views (Fig. 2). An advantage of ABUS is reconstructed coronal images showing “black holes” (Fig. 2), and “retraction phenomenon”, which has a high sensitivity and specificity for cancer (22). Our experience with the first generation ABUS was that “black holes” were occasionally not seen in small cancers (Fig. 3a). Improvements in technology may overcome this limitation. Comparison of HHUS and ABUS shows minor quality differences between the two techniques (Fig. 3a–c). However, studies comparing HHUS and ABUS have shown conflicting results. One study showed comparable diagnostic accuracy between the two techniques (23), whereas another study reported that only 57–78% of HHUS-detected cancers were identified with ABUS (12).

ABUS has two main weaknesses. First, shadowing from fibroglandular tissue can usually be solved with HHUS by applying compression and rotation of the transducer, but is a problem using ABUS. Shadowing was the reason for most missed cancers and most false positives in our study. Incomplete coverage is a further

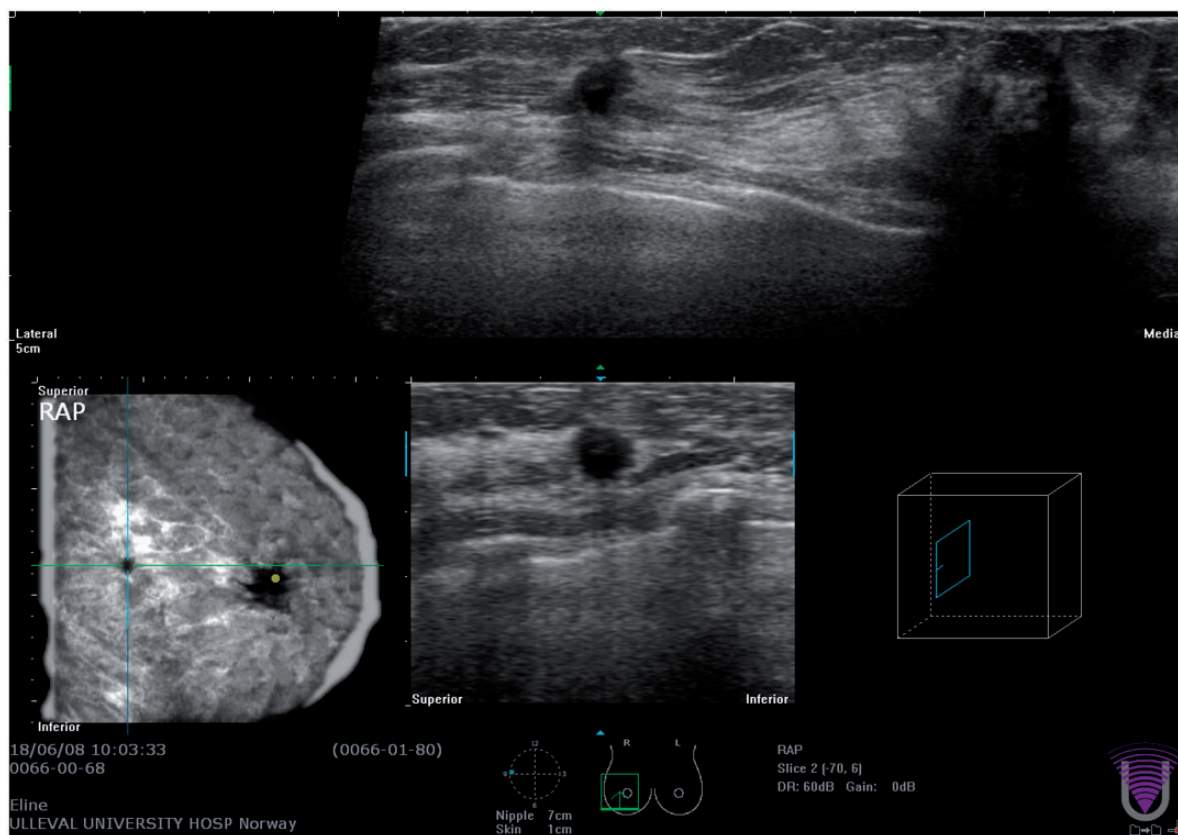


Fig. 2. Three-dimensional ABUS images shows a 15 mm invasive ductal carcinoma in the lateral part of the right breast on axial (transverse) view (top) and reconstructed coronal (bottom left) and sagittal (bottom right) views. The cancer is easily seen as a “black hole” (cross) on the coronal view. A yellow spot marks the location of the nipple on the coronal view.

reason for missing cancers using ABUS (24). Chang et al. reported that only 57–78% of HHUS-detected cancers were identified with ABUS (12). Radiologists can adjust focal zone, depth compensation gain, and transducer compressibility with HHUS whereas ABUS offer only fixed parameters (24). Second, a further limitation is the reported ABUS reading time of 5–10 min for a bilateral examination (12,16,23), which is in accordance with our study. This is too time-consuming for population-based screening.

Results on ABUS regarding interobserver variability have been conflicting. Very good agreement regarding BI-RADS classification of HHUS and ABUS (25) as well as only fair agreement has been reported (26). Agreement between mammography and ABUS was poor due to a high percentage of false positive ultrasound interpretations (27). Kappa values for ABUS alone was only fair in our study (Table 3), and the fair agreement for combined reading (Table 3) is in accordance with a recently published study (28). The inter-observer variability in our study needs some comments. First, the readers had different experience with ABUS. Second, the reading environment was different, and interruptions were not recorded. Finally, in this

“multi-institutional” study the readers might have applied different thresholds for BI-RADS categories.

Alternatives to ABUS in high-volume screening for women at normal risk with dense breasts include MRI and digital breast tomosynthesis (DBT). MRI could be carried out using “short first-pass technique” (11), but costs and availability are limiting factors. A promising modality for screening is DBT. Two studies have shown significantly higher cancer detection rates using tomosynthesis (29,30). DBT requires less manpower, and image acquisition and interpretation can be performed much faster than ABUS (30).

Our study has limitations. This was a retrospective study, and prospective trials would be mandatory before ABUS is implemented in screening. Second, the study population was small and did not permit stratified analysis on lesion types and breast density. Evaluation of sensitivity for small cancers would have been of most interest. A further limitation was that image interpretations were not carried out under equal and controlled reading environment. Finally, two readers had no experience with ABUS prior to the study and the pre-study training might have been insufficient.

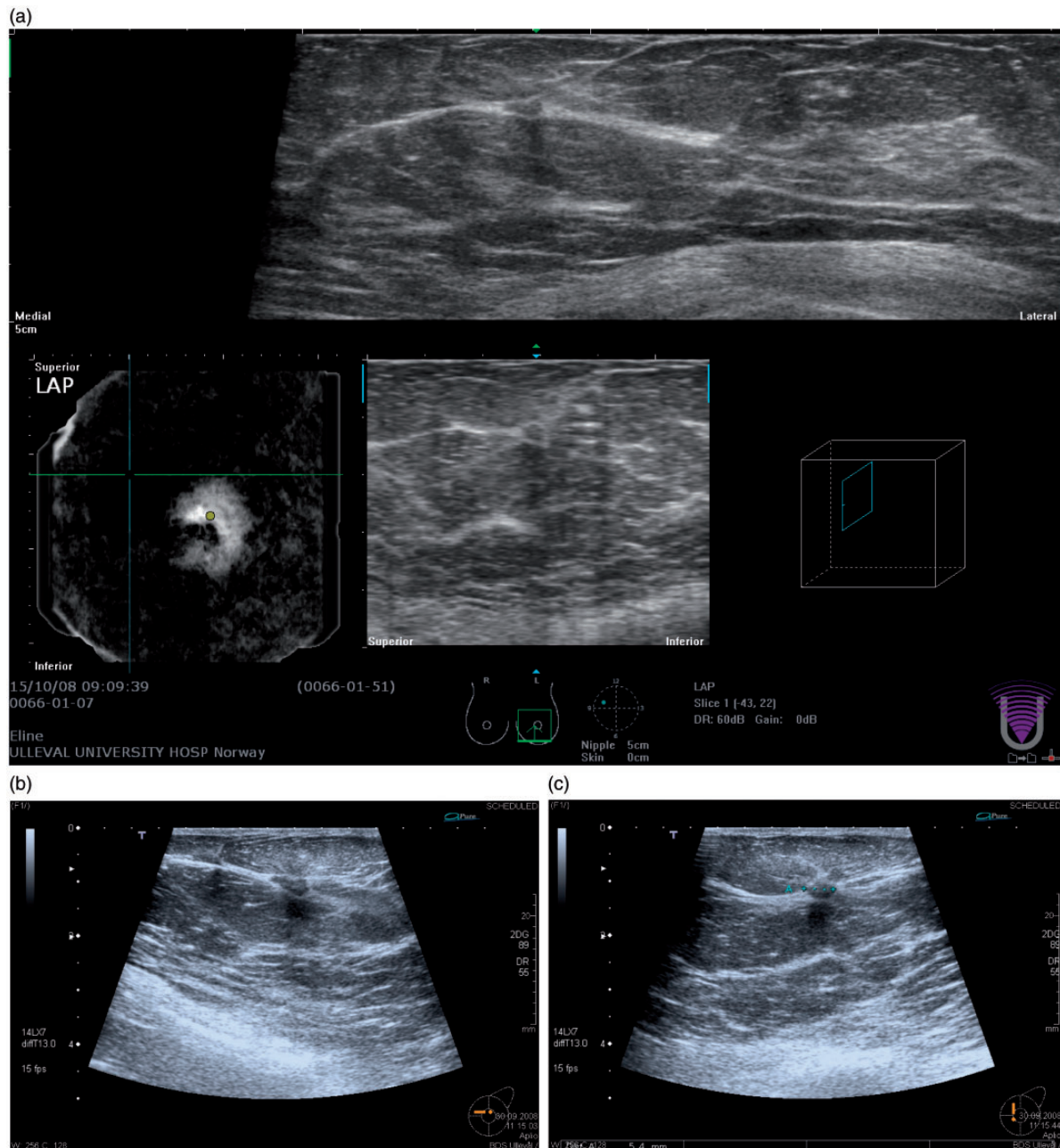


Fig. 3. (a) 3D ABUS multiplanar images shows a small screening-detected invasive ductal carcinoma (6 mm in diameter on histology) in the medial-upper quadrant of the left breast: Axial view (top), coronal view (bottom left), and sagittal image (bottom right). Note that no “black hole” is seen on the coronal image. A yellow spot marks the nipple. HHUS of the same patient shows the 6 mm invasive ductal carcinoma on (b) axial view (left) and (c) sagittal view (right). The HHUS images are somewhat zoomed, but otherwise the cancer is similarly presented on both techniques.

In conclusion, combined mammography and ABUS reading by the same radiologist improves diagnostic performance and results in higher observer agreement. Consequently, combined reading mode should be “standard” if ABUS is implemented in screening for women with dense breasts. Prospective studies are necessary before implementation of ABUS can be recommended in population-based screening.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Tabar L, Vitak B, Chen THH, et al. Swedish two-county trial: Impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* 2011;260:658–663.

2. Hofvind S, Ursin G, Tretli S, et al. Breast cancer mortality in participants of the Norwegian breast cancer screening program. *Cancer* 2013;119: 3106–3112.
3. Kuhl CK, Schrading S, Leutner CC, et al. Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. *J Clin Oncol* 2005;23:8469–8476.
4. Berg WA, Blume JD, Cormack JB, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA* 2008;299:2151–2163.
5. Buchberger W, Niehoff A, Obrist P, et al. Clinically and mammographically occult breast lesions: Detection and classification with high-resolution sonography. *Semin Ultrasound CT MR* 2000;21:325–336.
6. Corsetti V, Houssami N, Ferrari A, et al. Breast screening with ultrasound in women with mammography-negative dense breasts: Evidence on incremental cancer detection and false positives, and associated cost. *Eur J Cancer* 2008;44:539–544.
7. Crystal P, Strano SD, Shcharynski S, et al. Using sonography to screen women with mammographically dense breasts. *Am J Roentgenol* 2003;181:177–182.
8. Kaplan SS. Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue. *Radiology* 2001;221:641–649.
9. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: An analysis of 27,825 patient evaluations. *Radiology* 2002;225:165–175.
10. Nothacker M, Duda V, Hahn M, et al. Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. *BMC Cancer* 2009;9:335.
11. Fischer U, Korthauer A, Baum F, et al. Short first-pass MRI of the breast. *Acta Radiol* 2012;53:267–269.
12. Chang JM, Moon WK, Cho N, et al. Breast cancers initially detected by hand-held ultrasound: detection performance of radiologists using automated breast ultrasound data. *Acta Radiol* 2011;52:8–14.
13. Skaane P, Engedal K, Skjennald A. Interobserver variation in the interpretation of breast imaging: Comparison of mammography, ultrasonography, and both combined in the interpretation of palpable noncalcified breast masses. *Acta Radiol* 1997;38:497–502.
14. Hooley RJ, Greenberg KL, Stackhouse RM, et al. Screening US in patients with mammographically dense breasts: Initial experience with Connecticut Public Act 09-41. *Radiology* 2012;265:59–69.
15. Schaefer FKW, Waldmann A, Katalinic A, et al. Influence of additional breast ultrasound on cancer detection in a cohort study for quality assurance in breast diagnosis – analysis of 102,577 diagnostic procedures. *Eur Radiol* 2010;20:1085–1092.
16. Kelly KM, Dean J, Comulada WS, et al. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. *Eur Radiol* 2010;20:734–742.
17. Oberaigner W, Buchberger W, Frede T, et al. Introduction of organised mammography screening in Tyrol: results of a one-year pilot phase. *BMC Public Health* 2011;11:91.
18. Parris T, Wakefield D, Frimmer H. Real world performance of screening breast ultrasound following enactment of Connecticut Bill 458. *Breast J* 2013;19:64–70.
19. Skaane P, Engedal K. Analysis of sonographic features in the differentiation of fibroadenoma and invasive ductal carcinoma. *Am J Roentgenol* 1998;170:109–114.
20. 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.
21. Ciatto S, Bernardi D, Calabrese M, et al. A first evaluation of breast density assessment by Qantra software as compared to visual classification. *The Breast* 2012;21:503–506.
22. Zhang J, Lai X-J, Zhu Q-L, et al. Interobserver agreement for sonograms of breast lesions obtained by an automated breast volume scanner. *Eur J Radiol* 2012;81:2179–2183.
23. Wang H-Y, Jiang Y-X, Zhu Q-L, et al. Differentiation of benign and malignant breast lesions: A comparison between automatically generated breast volume scans and handheld ultrasound examinations. *Eur J Radiol* 2012;81:3190–3200.
24. Kim SH, Kang BJ, Choi BG, et al. Radiologists' performance for detecting lesions and the interobserver variability of automated whole breast ultrasound. *Korean J Radiol* 2013;14:154–163.
25. Wenkel E, Heckmann M, Heinrich M, et al. Automated breast ultrasound: Lesion detection and BI-RADS classification – a pilot study. *Fortschr Roentgenstr* 2008;180:804–808.
26. Wojcinski S, Farrokh A, Hille U, et al. The automated breast volume scanner (ABVS): initial experiences in lesion detection compared with conventional handheld B-mode ultrasound: a pilot study of 50 cases. *Int J Womens Health* 2011;3:337–346.
27. Stoblen F, Landt S, Stelkens-Gebhardt R, et al. First evaluation of the diagnostic accuracy of an automated 3D ultrasound system in a breast screening setting. *Anticancer Research* 2011;31:2569–2574.
28. Drukker K, Horsch KJ, Pesce LL, et al. Interreader scoring variability in an observer study using dual-modality imaging for breast cancer detection in women with dense breasts. *Acad Radiol* 2013;20:847–853.
29. Ciatto S, Houssami N, Bernardi D, et al. Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *Lancet Oncol* 2013;14:583–589.
30. Skaane P, Bandos AI, Gullien R, et al. Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. *Radiology* 2013;267:47–56.