의학박사 학위논문

Background Echotexture Classification using Automated Breast Ultrasound for Breast Cancer Screening: Interobserver Agreement and Screening Performance

검진을 목적으로 하는 자동 유방 초음파에서 배경 에코 분류: 관찰자간의 일치도 및 검진 진단능

울 산 대 학 교 대 학 원

의 학 과

방민서

Background Echotexture Classification using Automated Breast Ultrasound for Breast Cancer Screening: Interobserver Agreement and Screening Performance

지 도 교 수 원영철

이 논문을 의학박사 학위 논문으로 제출함.

2022 년 2 월

울 산 대 학 교 대 학 원
의 학 과
방민서

방민서의 의학박사 학위 논문을 인준함.

심사위원장	이 종 화	(인)
심사위원	최 혜 정	(인)
심사위원	원 영 철	(인)
심사위원	강 병 성	(인)
심사위원	김 성 헌	(인)

울 산 대 학 교 대 학 원

2022 년 2 월

ABSTRACT

Purpose: To prospectively evaluate the interobserver agreement for background echotexture assessments during screening with automated breast ultrasound (ABUS) and to assess the difference in the screening performance of ABUS, according to the background echotexture.

Materials and methods: We enrolled asymptomatic women aged 40–49 years from three participating centers between 2017 and 2019. Two radiologists at each center, specialized in breast imaging, classified background echotexture using a four-category classification (homogeneous 1, homogeneous 2, heterogeneous 1, and heterogeneous 2). The interobserver agreement was evaluated using kappa statistics. The recall rate, cancer yield, sensitivity, specificity, positive predictive value, and negative predictive value were calculated by dichotomizing the four categories into a two-category classification (homogeneous and heterogeneous) of background echotexture.

Results: A total of 990 women were included in the study. Almost perfect interobserver agreement (κ=0.825 and 0.812) was observed between the radiologists using the four- and two-category classifications. The recall rate was 8.2% in the homogeneous group and 13.0% in the heterogeneous group, with a significant difference ($p \leq 0.05$). The cancer yield was 2.89 per 1,000 screens in the homogeneous group and 10.06 per 1,000 screens in the heterogeneous group, with a no significant difference ($p>0.05$). Sensitivity, specificity, positive predictive value and negative predictive value were 100%, 91.9%, 3.5%, and 100% in the homogeneous group and 100%, 88.1%, 7.6%, and 100% in the heterogeneous group.

Conclusions: Background echotexture assessment had almost perfect interobserver agreement on screening with ABUS. The recall rate was significantly higher in women with a heterogeneous background echotexture.

Key words: Automated breast ultrasound, breast cancer, screening, background echotexture

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Mammography has long been the mainstay of breast imaging, owing to its application as a screening tool with a proven reduction in deaths caused by breast cancer [1]. However, the sensitivity of mammography for detecting breast cancer may vary depending on the mammographic density of the breast, which is composed of fat and fibroglandular tissue [2–4]. Among East Asian women, many women have high mammographic breast density, and in particular, women in their forties have the highest incidence of breast cancer, making them vulnerable to breast cancer diagnosis of mammography [5–7].

The wide variability in tissue composition seen on mammography can also be observed on ultrasound imaging, where it is defined as the background echotexture [8]. The heterogeneous background echotexture of the breast may affect the sensitivity of ultrasound for lesion detection, similar to the reduction in mammographic sensitivity for the detection of small lesions resulting from the increase in breast density [2–4,9].

Several studies have analyzed the background echotexture of breasts on ultrasonography [8–12]. Kim et al. [5] classified background echotexture into four categories. Their method demonstrated moderate interobserver agreement and proved that a heterogeneous background echotexture was associated with mammographically dense breast tissue. Ko et al. [10] reported that background echotexture, which was associated with background parenchymal enhancement on magnetic resonance imaging (MRI), may be a good predictor of background parenchymal enhancement. However, no study has investigated background echotexture using screening automated breast ultrasound (ABUS).

The advantages of the recently developed ABUS include the reduction in the need for supervision by radiologists during examination, acquisition of three-dimensional volume data, and convenience in obtaining a second reading; however, the nonavailability of specialized studies such as Doppler ultrasound or elastography and axilla evaluations is a disadvantage of ABUS [13,14]. There is a growing demand for breast ultrasonography screening, which can potentially be met by ABUS, instead of handheld ultrasound [13–15]. Over the last ten years, several studies have proved that supplemental screening with ABUS yielded a high diagnostic performance, similar to handheld US screening [16–19].

Three Korean academic medical centers conducted prospective study aimed to evaluate the diagnostic performance of screening ABUS in Korean women in their forties without mammography, and found that the diagnostic performance of screening ABUS was sufficient to be an alternative to screening mammography [20]. The present study is a part of that and focuses on background echotexture in screening ABUS. We prospectively assessed background echotexture using a newly proposed classification in screening ABUS and evaluated interobserver agreement for background echotexture. In addition, we investigated whether background echotexture could affect screening diagnostic performance.

MATERIALS AND METHODS

This study was approved by the institutional review boards of the participating centers. Written informed consent was obtained from all the participants. The Korean Society of Breast Imaging and Korean Society for Breast Screening (No. KSBI & KSFBS-2015-01) supported this study.

Study design

The three participating centers conducted this observer study on background echotexture at the time of participant enrollment. We enrolled 990 asymptomatic women who met the eligibility criteria from the three university hospitals between February 2017 and October 2019. We excluded women with a history of cancer, women who had undergone mammoplasty, and women who were currently pregnant or lactating. We collected demographic data and information on the risk factors of breast cancer at enrollment thorough an interview with a questionnaire.

Examination with ABUS

All ABUS examinations were conducted using the ACUSON S2000 ABVS (Siemens Medical Solutions, Mountain View, CA, USA) in combination with a 15-cm-wide linear array transducer with a 5–14 MHz bandwidth by experienced technicians. Three scans of volume data were obtained from each breast: anteroposterior volume, which covered the central part of the breast; medial volume, which covered the inner and inferior parts; and lateral volume, which covered the upper and outer parts of the breast. Additional views were acquired to cover the entire breast tissue in participants with larger breasts. The volume images were automatically transferred to a dedicated workstation.

Image interpretation

Images were interpreted through double reading by two radiologists at each hospital. They had 9 to 17 years of experience with breast imaging. These two radiologists had a consensus reading meeting in case of disagreements regarding the assessment of the category and background echotexture. Multiplanar images in three different planes (axial, sagittal, and coronal) were used during the interpretation. The radiologists were blinded to clinical information of the participants.

Before the study commenced, the participating radiologists established interpretation criteria for ABUS screening by modifying the Breast Imaging Reporting and Data System (BI-RADS) and Japanese association of breast and thyroid sonology classification (Table 1) [13,21,22]. The participating radiologists' assessments were classified into one of the following categories: 0, incomplete; 1, negative; 2, benign; 3, probably benign; 4, suspicious; and 5, highly suggestive of malignancy.

The radiologists assessed the background echotexture, based on the interpretation criteria they established, at the most heterogeneous part of the entire breast parenchyma, 2–3 cm above the nipple or upper outer quadrant, based on the visually estimated proportion and uniformity of the isoechoic or hypoechoic areas that consisted of fat lobules, the terminal ductolobular unit, and normal mammary ducts in the fibroglandular tissue. The imaged tissue was classified into one of the four background echotexture categories. Homogeneous echotexture 1 (E1) was defined as breast parenchyma that appeared mostly homogeneously hyperechoic, with minimal isoechoic or hypoechoic areas. Homogeneous echotexture 2 (E2) was defined as hyperechoic parenchyma with uniform appearance of the hypoechoic ducts. Heterogeneous echotexture 1 (E3) was defined as parenchyma with echogenic fibroglandular tissue interspersed with isoechoic fat or irregular hypoechoic ducts with greater dilation, which may affect sensitivity. Heterogeneous echotexture 2 (E4) was defined as fibroglandular tissue showing isoechoic or hypoechoic areas, which may mimic a mass. A standard set of reference images of the background echotexture of the breast is illustrated in Figure 1.

Reference standard

We recalled and reexamined all the women with suspicious findings on ABUS with a handheld US scanner. The procedure was performed by a breast imaging specialized radiologist using EPIQ 7 or IU22 (Philips, Bothell, WA, USA). US-guided 14G or 16G core-needle biopsy was carried out for suspicious lesions, if needed. The results of the pathologic examination were considered as the reference standard.

Statistical analyses

In this study, the observer used a four category classification method when analyzing the background echotexture, but in statistical analysis, it were classified into two category again and analyzed.

The independent t-test, Pearson's chi-squared test, and one-way analysis of variance were used to compare the mean background echotexture based on the risk factors for breast cancer.

The radiologists' agreement on the background echotexture was assessed using kappa statistics. We used the following definitions to interpret the kappa coefficients (κ): κ-values less than 0.20 indicated poor agreement; κ-values of 0.21–0.40 indicated fair agreement; κ-values of 0.41–0.60 indicated moderate agreement; κ-values of 0.61–0.80 indicated substantial agreement; and κ-values of 0.81–1.00 indicated almost perfect agreement [23].

The results were considered positive for category 0, 3, or higher if further evaluations were needed. A category 1 or 2 was considered negative. We set the performance measures as follows: recall rate, or the percentage of screening examinations with positive results (categories 0, 3, 4, and 5); cancer yield, or the number of true positive screens per 1000 screens; positive predictive value, or the percentage of screening examinations with positive results (categories 0, 3, 4, and 5) resulting in a tissue diagnosis of cancer; and negative predictive value, or the percentage of screening examinations with negative results (categories 1 and 2) resulting in a true negative final diagnosis. Fisher's exact test and Pearson's chi-squared test were used to evaluate the diagnostic values within the groups.

Statistical analyses were performed using SPSS software (version 21.0; IBM Corp., Armonk, NY), and P-values≤0.05 were considered statistically significant.

RESULTS

The background echotexture of 990 patients was classified as homogeneous echotexture 1 (E1) in 411 assessments (41.5%), homogeneous echotexture 2 (E2) in 280 (28.3%)

assessments, heterogeneous echotexture 1 (E3) in 236 (23.8%) assessments, and heterogeneous echotexture 2 (E4) in 63 (6.4%) assessments.

The distribution of the background echotexture according to the risk factors for breast cancer is summarized in Table 2. No statistically significant difference was identified in the background echotexture according to age, body mass index (BMI), age at menarche, menopausal status, family history of breast cancer, and breast feeding. The four background echotexture categories were dichotomized into homogeneous (E1, E2) and heterogeneous (E3, E4). The dichotomized background echotexture classes did not show any statistical significance with respect to the risk factors for breast cancer (Table 3).

Interobserver agreement for background echotexture

The interobserver agreement values for background echotexture assessment using the four-category classification indicated an almost perfect agreement (κ-value=0.825, p=0.000) between the two radiologists. The agreement was almost perfect (κ-value=0.812, p=0.000) even after the dichotomization of the categories into homogeneous (E1, E2) and heterogeneous (E3, E4) (Table 4). One hundred discrepancies were identified in 990 cases using the four-category classification; the number of discrepancies was the highest with the E2 and E3 categories.

Screening performance of ABUS according to background echotexture

Of the 990 assessments, 894 (90.3%) cases were initially classified as category 1 or 2 and the 96 (9.7%) cases were classified into category 3, 4, and 5, which were recalled. Categories 3, 4, and 5 consisted of 71, 22, and 3 assessments, respectively. The recall rate was 8.2% for women with homogeneous echotexture, but 13.0% for women with heterogeneous echotexture and this difference was statistically significant (p=0.018) (Table 5).

The overall cancer yield was 5.1 cancers per 1000 screens. Two cancers occurred in the homogeneous echotexture group and three cancers occurred in the heterogeneous echotexture group (Figure 2,3). The cancer yield was 2.89 per 1000 screens for women with homogeneous echotexture, but 10.06 per 1000 screens for women with heterogeneous echotexture; however, this difference was statistically not significant (p=0.164) (Table 5).

The specificity was 91.9% and 88.1% in the homogeneous and heterogeneous echotexture groups, respectively, while the sensitivity was 100% in both groups. The positive predictive value was 3.5% and 7.6% in the homogeneous and heterogeneous echotexture groups, respectively, while the negative predictive value was 100% in both groups (Table 5).

DISCUSSION

Although the 5th edition of the American College of Radiology Breast Imaging Reporting and Data System (ACR BI-RADS) classified background echotexture into homogeneous-fat, homogeneous-fibroglandular, and heterogeneous, it demonstrated limitations in reflecting the various background echotextures of the breast in Asian women [22]. The proportion of breasts with homogeneous-fat echotexture is low among Asian women [8,9]. Moreover, the heterogeneous echotexture category of the ACR BI-RADS classification comprises a wide spectrum of background echotextures [22]. Two studies by Kim et al. [8,11] modified the four-category classifications of background echotexture by creating the homogeneous, mild heterogeneous, moderate heterogeneous and marked heterogeneous categories; they emphasized on the proportion of fibroglandular tissue in the entire breast in 2013 and that of the isoechoic or hypoechoic areas in fibroglandular tissue in 2017. We also established a four-category echotexture classification in our study, with greater emphasis on the uniformity and proportion of isoechoic or hypoechoic areas: homogeneous 1 (E1), homogeneous 2 (E2), heterogeneous 1 (E3), and heterogeneous 2 (E4).

Lobule, duct and dense interlobular stromal fibrous tissues appear as isoechoic or hypoechoic structures, whereas loose stromal fibrous connective tissue appears hyperechoic on ultrasonography. The differences between the echogenicity of the ductal systems and stromal fibrous tissues may result in a heterogeneous background echotexture [24]. Ko et al. [9] evaluated the contralateral breast in 140 patients with breast cancer and reported that younger age and premenopause were associated with heterogeneous echotexture. Kim et al. [8] also reported that younger age and premenopause were associated with heterogeneous echotexture in addition to a nulliparous status and lower BMI. However, in our study, background echotexture did not differ significantly depending on age, menopausal status, and BMI. Previous studies included participants of various ages, but only women in their forties were included in our study; therefore, the deviation between the participants' age and hormonal status was not significant, which could have influenced the results of our study [8,9].

Background echotexture is determined by the interpreter's estimation of the proportion of isoechoic or hypoechoic areas. Thus, there is an inherent degree of subjectivity in the determination of background echotexture on breast ultrasonography. Kim et al. [8,11] demonstrated above-moderate interobserver agreement among eight radiologists (κ-value=0.67) in 2013 and among eleven radiologists (κ-value=0.45) in 2017 using the four-category background echotexture classification: homogeneous, mild heterogeneous, moderate heterogeneous, and marked heterogeneous. Berg et al. [12] reported a κ-value of 0.30 for the heterogeneity of background echotexture among eleven radiologists using three categories: homogeneous, focal heterogeneous, and diffuse heterogeneous (Table 6). The interobserver agreement values (κ-value=0.827 and 0.816) in our study were higher than those reported in previous studies [8,11,12]. Although we evaluated interobserver agreement between two radiologists, our study population consisted of a larger number of healthy women. We conducted consensus meetings regarding the assessment of background echotexture before the study and had already completed assessment training. Moreover, we used the reference images for each background echotexture assessment. Thus, complete nonselective documentation of the image data, examiner independence, and reproducibility of ABUS may also lower the interobserver variability of background echotexture assessment [13,25].

The present study showed that the recall rate of the heterogeneous-echotexture group was higher than that of the homogeneous echotexture group, similar to the higher recall rate and lower sensitivity associated with higher mammographic breast density. Several studies reported a highly variable recall rate (2.2%–28.4%) with supplemental ABUS in mammographically screened women with dense breasts [26–28]. We conducted independent ABUS for all participants at enrollment time in our study, irrespective of mammographic density, and found a recall rate of 9.6%, irrespective of the background echotexture. It may be difficult to identify pathologic lesions and differentiate them from normal parenchymal tissue using ultrasound in women with heterogeneous background echotextures. The frequently visible artifacts of ABUS that are also important factors that increase recall rates in women with heterogeneous background echotexture include marked shadowing due to fibrotic breasts, poor contact, and the nipple [13,14,25].

Although it was not statistically significant, our ABUS examination results revealed a higher cancer yield in the heterogeneous echotexture group than in the homogeneous echotexture group. Several studies reported that the breast cancer risk was four to six times higher in patients with extremely high density than that in patients with lower density [28–30]. There has been no study on the associattion between background echotexture of ultrasound exam and breast cancer, but a report that studied it was recently published. This study found that the amount of glandular tissue shown as isoecho or hypoecho was independently associated with breast cancer[31]. Since heterogeneous background echotexture contain more isoechoic or hypoechoic portion, the breast cancer risk of women with them may be higher than that of women with homogeneous background echotexture.

This study had some limitations. First, the study population included only women in their forties, which is the age of perimenopause. Thus, it was difficult to evaluate the effects of age and menopause status on ultrasonographic background echotexture. Thus, some of the results may not be generalizable to other age groups. Second, this study was solely based on examinations at the initial enrollment of the participants and did not include follow-up results. Verification of the diagnostic performance using follow-up results is needed in the future. Third, we included only two readers in each hospital on the interobserver agreement analysis of background echotexture assessment. Thus, further study are needed to analyze the interobserver agreement between a larger number of readers. Forth, we used a new background echotexture classification. In addition, this is the only study using ABUS for background echotexture. There may be limitations in the validity of direct comparisons with the results of previous studies that used other background echotexture classifications.

CONCLUSION

This study demonstrated that background echotexture provided a reliable criterion for classification, with a relatively high interobserver agreement. Moreover, background echotexture affected the screening performance of ABUS. Therefore, background echotexture should be considered while reading the screening ABUS, similar to the inclusion of breast composition in reading the screening mammography.

REFERENCES

1. Niell BL, Freer PE, Weinfurtner RJ, Arleo EK, Drukteinis JS. Screening for breast cancer. Radiol Clin North Am 2017;55:1145–1162.

2. Mandelson MT, Oestreicher N, Porter PL, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92:1081–1087.

3. Buist DS, Porter PL, Lehman C, Taplin SH, White E. Factors contributing to mammography failure in women aged 40-49 years. J Natl Cancer Inst 2004;96:1432–1440.

4. Porter GJ, Evans AJ, Cornford EJ, et al. Influence of mammographic parenchymal pattern in screening detected and interval invasive breast cancers on pathologic features, mammographic features, and patient survival. AJR Am J Roentgenol 2007;188:676–683.

5. Yap YS, Lu YS, Tamura K et al. Insights into breast cancerin the East vs the West: a review. JAMA Oncol 2019;5:1489–1496.

6. Rajaram N, Mariapun S, Eriksson M et al. Differences in mammographic density between Asian and Caucasian populations: a comparative analysis. Breast Cancer Res Treat 2017; 161:353–362.

7. Leong SP, Shen ZZ, Liu TJ et al.Is breast cancer the same disease in Asian and Western

countries? World J Surg 2010;34:2308–2324.

8. Kim WH, Lee SH, Chang JM, Cho N, Moon WK. Background echotexture classification in breast ultrasound: inter-observer agreement study. Acta Radiol 2017;58:1427–1433.

9. Ko KH, Son EJ, Kim IW. Accuracy of ultrasound for preoperative assessment of tumor size in patients with newly diagnosed breast cancer: is it affected by the background parenchymal echotexture? J Ultrasound Med 2018;37:2621–2630.

10. Ko ES, Lee BH, Choi HY, Kim RB, Noh WC. Background enhancement in breast MR: correlation with breast density in mammography and background echotexture in ultrasound. Eur J Radiol 2011;80:719–723.

11. Kim WH, Moon WK, Kim SJ, et al. Ultrasonographic assessment of breast density. Breast Cancer Res Treat 2013;138:851–859.

12. Berg WA, Blume JD, Cormack JB, Mendelson EB. Operator dependence of physician-performed whole-breast US: lesion detection and characterization. Radiology 2006;241:355–365.

13. Kim SH, Kim HH, Moon WK. Automated breast ultrasound screening for dense breasts. Korean J Radiol 2020;21:15–24.

14. Rella R, Belli P, Giuliani M, et al. Automated breast ultrasonography (ABUS) in the

screening and diagnostic setting: indications and practical use. Acad Radiol 2018;25:1457–1470.

15. Aripoli A, Fountain K, Winblad O, et al. Supplemental screening with automated breast ultrasound in women with dense breasts: comparing notification methods and screening behaviors. AJR Am J Roentgenol 2018;210:W22–W28.

16. Schmachtenberg C, Fischer T, Hamm B, Bick U. Diagnostic performance of automated breast volume scanning (ABVS) compared to handheld ultrasonography with breast MRI as the gold standard. Acad Radiol 2017;24:954–961.

17. Yun G, Kim SM, Yun B, Ahn HS, Jang M. Reliability of automated versus handheld breast ultrasound examinations of suspicious breast masses. Ultrasonography 2019;38:264–271.

18. An YY, Kim SH, Kang BJ. The image quality and lesion characterization of breast using automated whole-breast ultrasound: a comparison with handheld ultrasound. Eur J Radiol 2015;84:1232–1235.

19. Meng Z, Chen C, Zhu Y, et al. Diagnostic performance of the automated breast volume scanner: a systematic review of inter-rater reliability/agreement and meta-analysis of diagnostic accuracy for differentiating benign and malignant breast lesions. Eur Radiol 2015;25:3638–3647.

20. Choi WJ, Kim SH, Shin HJ, et al. Automated breast US as the primary screening test for breast cancer among East Asian women aged 40–49 years: a multicenter prospective study. European Radiology 2021;31:7771–7782

21. Ohuchi N, Suzuki A, Sobue T, et al. Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomised controlled trial. Lancet 2016;387:341–348.

22. D'Orsi CJ, Sickles EA, Mendelson EB, et al. ACR BI-RADS® Atlas, Breast Imaging Reporting and Data Syste, 5th ed. American College of Radiology, 2013.

23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.

24. Izumori A, Horii R, Akiyama F, Iwase T. Proposal of a novel method for observing the breast by high-resolution ultrasound imaging: understanding the normal breast structure and its application in an observational method for detecting deviations. Breast Cancer 2013;20:83–91.

25. Kim SH. Image quality and artifacts in automated breast ultrasonography. Ultrasonography 2019;38:83–91.

26. Wilczek B, Wilczek HE, Rasouliyan L, Leifland K. Adding 3D automated breast

ultrasound to mammography screening in women with heterogeneously and extremely dense breasts: report from a hospital-based, high-volume, single-center breast cancer screening program. Eur J Radiol 2016;85:1554–1563.

27. Giuliano V, Giuliano C. Improved breast cancer detection in asymptomatic women using 3D-automated breast ultrasound in mammographically dense breasts. Clin Imaging 2013;37:480–486.
28. Boyd NF, Byng JW, Jong RA, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. J Natl Cancer Inst 1995;87:670–675.
29. Byrne C, Schairer C, Wolfe J, et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst 1995;87:1622–1629.
30. Yaghjyan L, Colditz GA, Rosner B, Tamimi RM. Mammographic breast density and breast cancer risk: interactions of percent density, absolute dense, and non-dense areas with breast cancer risk factors. Breast Cancer Res Treat 2015;150:181–189.
31. Lee SH, Ryu HS, Jang MJ, et al. Glandular Tissue Component and Breast Cancer Risk in Mammographically Dense Breasts at Screening Breast US. Radiology 2021;301:57-65.
32. American College of Radiology. ACR BI-RADS® Atlas, Breast Imaging Reporting and

Data System, 4th ed. American College of Radiology, 2003

국문요약

목적: 검진 목적으로 시행한 자동 유방 초음파 검사에서 배경 에코를 분류하고 검사자간의 일치도를 측정하며 배경 에코에 따른 검진 진단능의 차이를 알아 본다.

대상 및 방법: 2017년부터 2019년까지, 3개의 병원에서 40대 무증상 여성을 모집하여 자동 유방 초음파 검사를 시행하였다. 각 병원의 두 명의 유방 영상을 전공한 영상의학과 의사가 4가지 배경 에코 분류법 (균질1, 균질2, 불균질1, 불균질2) 을 이용하여 자동 유방 초음파 영상의 배경 에코를 분류하였고 kappa 통계를 이용하여 두 명의 영상의학과 의사 사이의 일치도를 측정하였다. 4가지 배경 에코 분류법을 다시 2가지 배경 에코 분류법 (균질, 불균질)으로 이분화하여 배경 에코에 따른 소환율, 암 발생률, 민감도, 특이도, 양성예측률, 음성예측률을 계산해보았다.

결과: 이 연구에 총 990명의 여성이 참가하였다. 두 명의 영상의학과 의사사이에 4가지 배경 에코 분류법과 2가지 배경 에코 분류법을 이용하여 배경 에코 분석 일치도를 구했을 때 거의 완전 일치도 (κ=0.825 and 0.812) 를 보였다. 소환율은 균질한 배경 에코를 가진 여성 군은 8.2%, 불균질한 배경 에코를 가진 여성 군은 13.0%로 통계적으로 유의한 차이가 있었다 ($p \leq 0.05$). 암 발생률은 균질한 배경

에코를 가진 여성군은 1000명당 2.89개, 불균질한 배경 에코를 가진 여성군은 1000명당 10.06개로 통계적으로 유의하지 않았다 (p>0.05). 민감도, 특이도, 양성예측률 그리고 음성예측률은 균질한 배경 에코를 가진 여성군은 100%, 91.9%, 3.5%, 100%, 불균질한 배경 에코를 가진 여성군은 100%, 88.1%, 7.6%, and 100% 였다.

결론: 자동 유방 초음파에서 배경 에코 분류는 거의 완전 일치도를 보였다. 그리고 소환율은 불균질한 배경 에코를 가진 여성에서 유의하게 더 높은 것을 알 수 있었다.

Table 1. Category interpretation criteria for automated breast ultrasound screening

Finding	Category	Management
Simple cyst, intramammary lymph node, calcified fibroadenoma, fat-containing lesion Multiple, oval, circumscribed complicated cysts or masses Non-simple cysts in the setting of multiple or bilateral cysts (at least three cysts, with at least one in each breast) Round, circumscribed, solid mass (≤5 mm) Oval, circumscribed, parallel, solid mass (≤10 mm)	2	Follow-up after 1 year
Isolated complicated cyst Round, circumscribed, solid mass (>5 mm)	3	Follow-up after 6 months

Oval, circumscribed, parallel mass (>10 mm)		
Clustered microcysts		
Fat necrosis		
Well-defined intraductal lesion		
Others	4	Pathologic confirmation
Irregular, spiculated mass	5	Pathologic confirmation

Table 2. Background echotexture according to risk factors in enrolled women using four categories

		Homogeneous1 (E1)	**Homogeneous2 (E2)**	**Heterogeneous1 (E3)**	**Heterogeneous2 (E4)**	**p-value**
		n=411	n=280	n=236	n=63	
		Mean±SD	Mean±SD	Mean±SD	Mean±SD	
Age, years		43.93±3.12	43.75±3.08	43.96±3.06	43.59±3.01	0.720
BMI, kg/m^2		22.50±3.08	22.56±3.05	22.48±2.86	21.78±3.39	0.322
Menarche age, years		14.24±1.37	14.04±1.37	14.12±1.16	14.29±1.28	0.200
		N (%)	**N (%)**	**N (%)**	**N (%)**	**p-value**
Menopausal status	Premenopausal	339 (82.5)	231 (82.5)	203 (86.0)	54 (85.7)	0.667
	Perimenopausal	38 (9.2)	20 (7.1)	18 (7.6)	4 (6.3)	
	Postmenopausal	34 (8.3)	29 (10.4)	15 (6.4)	5 (7.9)	
Age at first birth, years	NA	56 (13.6)	49 (17.5)	40 (16.9)	11 (17.5)	0.430
	10–19	0 (0.0)	2 (0.7)	0 (0.0)	0 (0.0)	
	20–29	207 (50.4)	123 (43.9)	109 (46.2)	23 (36.5)	
	30–39	139 (33.8)	99 (35.4)	84 (35.6)	28 (44.4)	

	40–49	9 (2.2)	7 (2.5)	3 (1.3)	1 (1.6)	
Breast-feeding	Yes	286 (69.6)	192 (68.6)	163 (69.1)	43 (68.3)	0.991
	No	125 (30.4)	88 (31.4)	73 (30.9)	20 (31.7)	
Family history of breast cancer	Yes	29 (7.1)	34 (12.2)	23 (9.9)	9 (14.5)	0.078
	No	379 (92.9)	244 (87.8)	210 (90.1)	53 (85.5)	

BMI: body mass index, NA: not applicable, SD: standard deviation

Table 3. Background echotexture according to risk factors in enrolled women using dichotomized two categories

		Homogeneous (E1, 2)	Heterogeneous (E3, 4)	p-value
		n=691	n=299	
		Mean±SD	Mean±SD	
Age, years		43.85±3.10	43.90±3.05	0.845
BMI, kg/m^2		22.52±3.06	22.34±2.99	0.390
Menarche age, years		14.16±1.37	14.16±1.19	0.997
		N (%)	**N (%)**	**p-value**
Menopausal status	Premenopausal	571 (82.5)	256 (85.9)	0.374
	Perimenopausal	58 (8.4)	22 (7.4)	
	Postmenopausal	63 (9.1)	20 (6.7)	
Age at first birth, years	NA	105 (15.2)	51 (17.1)	0.513
	10–19	2 (0.3)	0 (0.0)	
	20–29	330 (47.7)	132 (44.3)	
	30–39	239 (34.5)	111 (37.2)	
	40–49	16 (2.3)	4 (1.3)	

Breast-feeding	Yes	479 (69.2)	205 (68.8)	0.894
	No	213 (30.8)	93 (31.2)	
Family history of breast cancer	Yes	63 (9.2)	32 (10.9)	0.406
	No	624 (90.8)	262 (89.1)	

BMI: body mass index, NA: not applicable, SD: standard deviation

Table 4. Interobserver agreement for background echotexture

	κ-value	p-value
Four-categories (E1 vs E2 vs E3 vs E4)	0.825	0.000
Two-categories (E1, 2 vs E3, 4)	0.812	0.000

$\kappa \leq 0.00$ Poor agreement, $0.00 \leq \kappa \leq 0.20$ Slight agreement, $0.21 \leq \kappa \leq 0.40$ Fair agreement, $0.41 \leq \kappa \leq 0.60$ Moderte agreement, $0.61 \leq \kappa \leq 0.80$ Substantial agreement, $\kappa > 0.80$ Almost perfect agreement

Table 5. Diagnostic performance of automated breast ultrasound for breast cancer screening stratified according to background echotexture

		Background echotexture		
		Homogeneous (E1, 2)	**Heterogeneous (E3, 4)**	**p-value**
Recall rate	N	57/692	39/298	0.018
	%	8.2%	13.0%	
Cancer yield	N	2/692	3/298	0.164
	per 1000 screens	2.89	10.06	
Sensitivity	N	2/2	3/3	
	%	100%	100%	
Specificity	N	635/690	259/295	
	%	91.90%	88.10%	
PPV	N	2/57	3/39	
	%	3.50%	7.60%	
NPV	N	635/635	259/259	
	%	100%	100%	

NPV: negative predictive value, PPV: positive predictive value

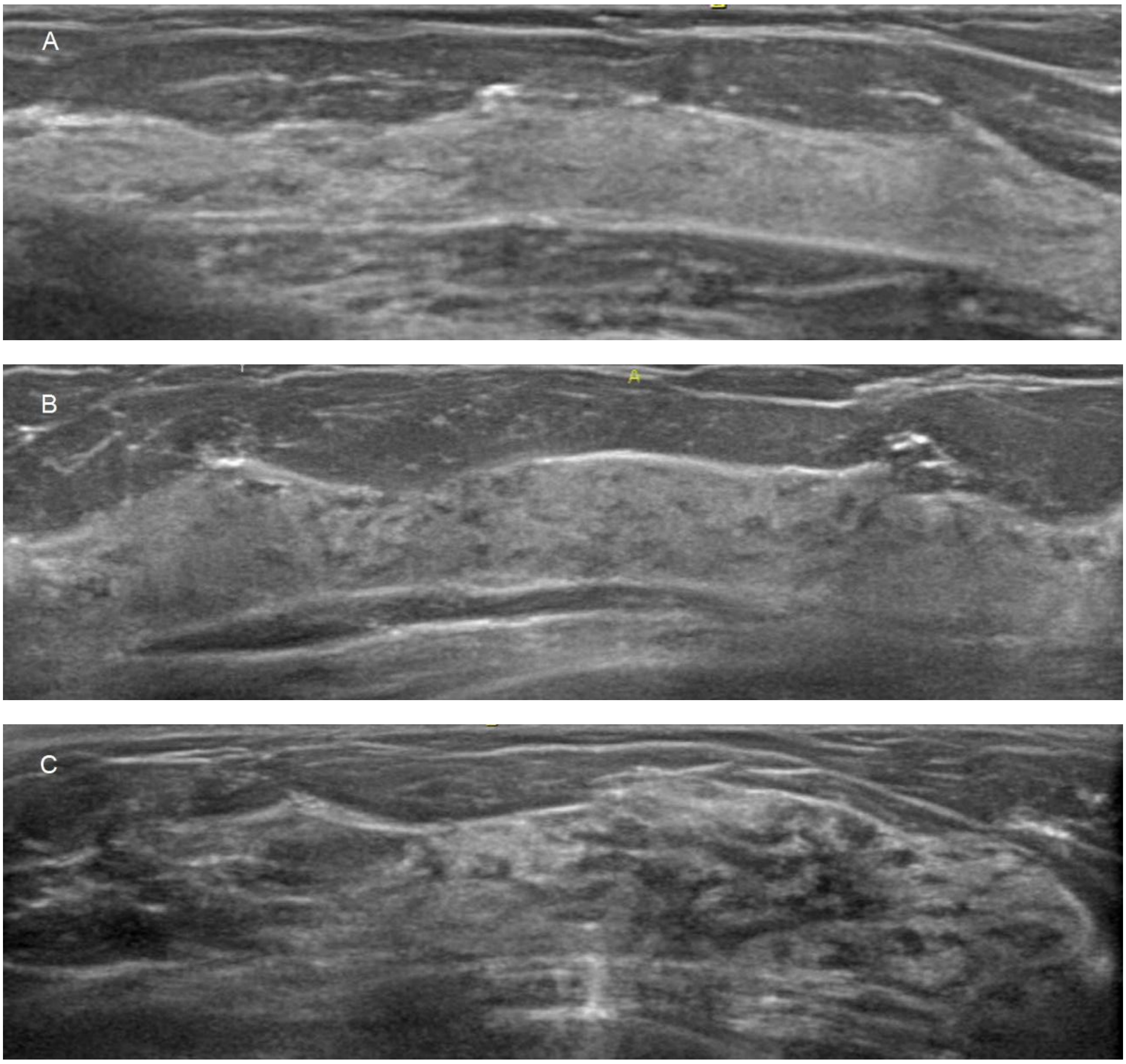
Table 6. Interobserver agreement studies about background echotexture of breast ultrasound.

	Present study	**Berg et al.[12]**	**Kim et al.[11]**	**Kim et al.[8]**
Classification	Uniformity and proportion of isoechoic or hypoechoic areas in fibroglandular tissue	ACR BI-RADS, 2003[32]	Proportion of fibroglandular tissue in breast	Proportion of isoechoic or hypoechoic areas in fibroglandular tissue
	Homogeneous 1 Homogeneous 2 Heterogeneous 1 Heterogeneous 2	Homogeneous Focal heterogeneous Diffuse heterogeneous	Homogeneous Mild heterogeneous Moderate heterogeneous Marked heterogeneous	Homogeneous Mild heterogeneous Moderate heterogeneous Marked heterogeneous
Study design	Prospective	Prospective	Prospective	Prospective
Study population	990 healthy women	10 patients with known breast	41 healthy women	38 healthy women

		lesions		
Nationality	South Korea	USA	South Korea	South Korea
Participating radiologists	6 (2 at each hospital)	11	8	13
κ-value	0.825	0.30	0.63	0.45

ACR BI-RADS: American College of Radiology Breast Imaging Data and Reporting System

Figure 1. Representative images of background echotexture of automated breast ultrasonography. (A) Homogeneous echotexture 1 (E1), (B) Homogeneous echotexture 2 (E2), (C) Heterogeneous echotexture 1 (E3), and (D) Heterogeneous echotexture 2 (E4).



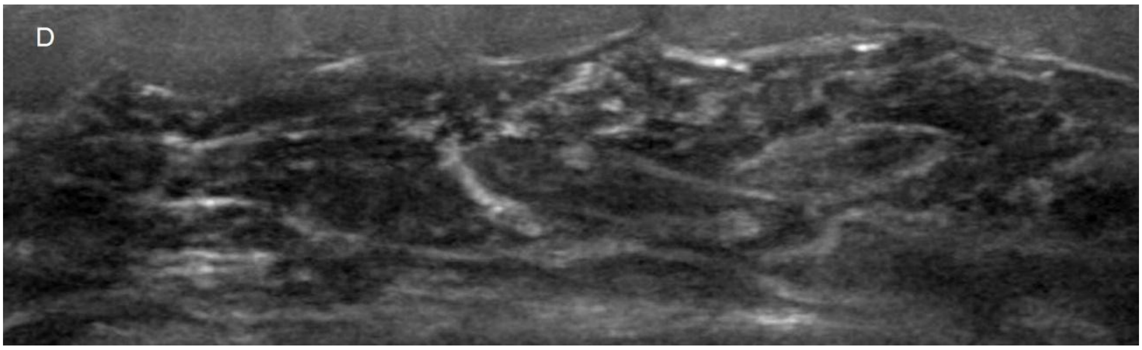
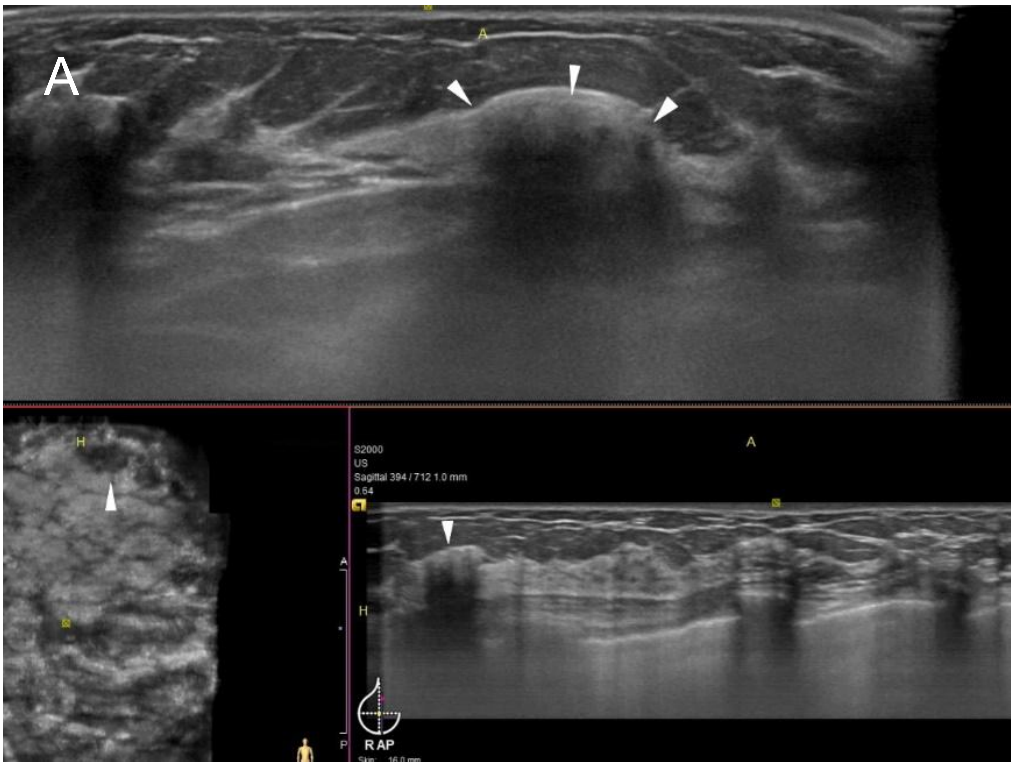

D

Figure 2. Cancer detected by ABUS screening in a 40-year-old woman with homogeneous echotexture. (A) Right anteroposterior (AP) axial view of ABUS shows homogeneous 2 (E2) background echotexture. Axial (upper column), coronal (right lower column), and sagittal (left lower column) ABUS images show a heterogeneous hypoechoic mass with indistinct margins (arrowheads) in the 12 o'clock direction. (B) Handheld US shows a mass with suspicious for malignancy in the 12 o'clock direction (arrowheads). Ultrasound-guided core biopsy was performed, and histopathologic examination revealed an invasive ductal carcinoma.



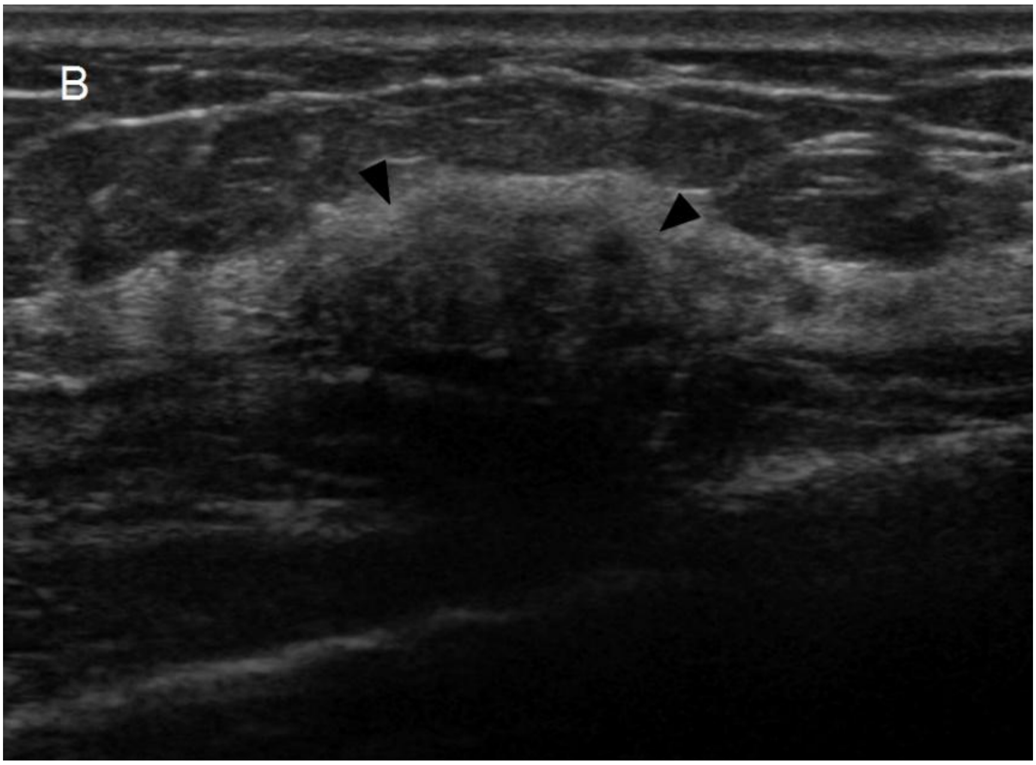
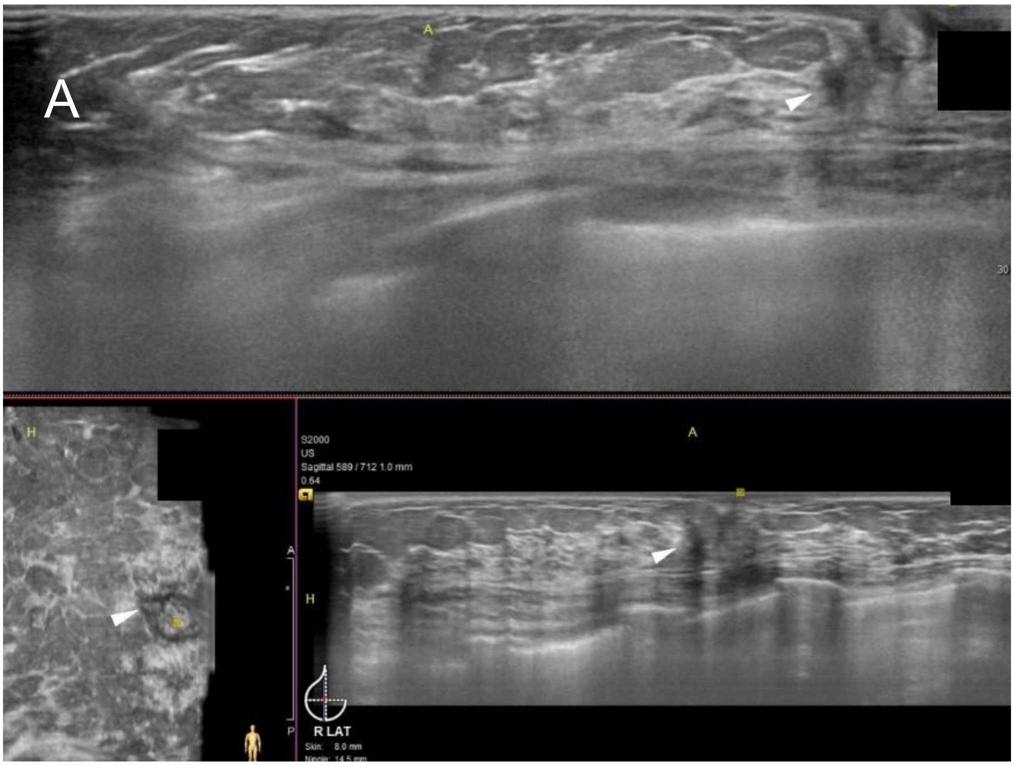
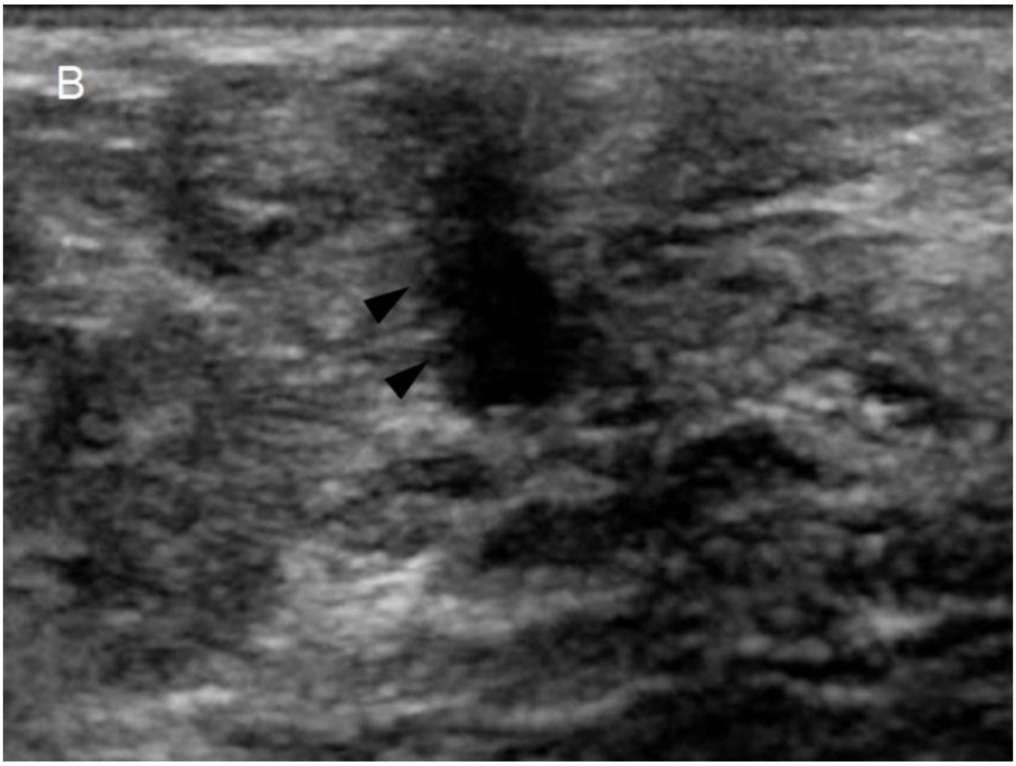

B

Figure 3. Cancer detected by ABUS screening in a 49-year-old woman with heterogeneous echotexture. Right lateral view of ABUS shows heterogeneous 1 (E3) background echotexture. Axial (upper column), coronal (right lower column), and sagittal (left lower column) images show an irregular spiculated hypoechoic mass (arrowheads) in the subareolar area. (B) Handheld US shows a mass with suspicious for malignancy in the subareolar area (arrowheads).Ultrasound-guided core biopsy was performed, and histopathologic examination revealed a mixed invasive ductal and mucinous carcinoma.




B