



Artificial intelligence in breast ultrasound: application in clinical practice

ULTRASONOGRAPHY

Hila Fruchtmann Brot, Victoria L. Mango

Memorial Sloan Kettering Cancer Center, New York, NY, USA

Ultrasound (US) is a widely accessible and extensively used tool for breast imaging. It is commonly used as an additional screening tool, especially for women with dense breast tissue. Advances in artificial intelligence (AI) have led to the development of various AI systems that assist radiologists in identifying and diagnosing breast lesions using US. This article provides an overview of the background and supporting evidence for the use of AI in hand held breast US. It discusses the impact of AI on clinical workflow, covering breast cancer detection, diagnosis, prediction of molecular subtypes, evaluation of axillary lymph node status, and response to neoadjuvant chemotherapy. Additionally, the article highlights the potential significance of AI in breast US for low and middle income countries.

Keywords: Artificial intelligence; Breast neoplasms; Computer-aided detection; Computer-aided diagnosis; Ultrasound

Key points: Artificial intelligence (AI) based detection and diagnostic decision support tools have the potential to serve an important clinical role in handheld breast ultrasound. More prospective studies are needed to understand the impact of AI on actual clinical diagnostic performance and how to incorporate AI into real-world clinical settings.

REVIEW ARTICLE

<https://doi.org/10.14366/usg.23116>
eISSN: 2288-5943
Ultrasonography 2024;43:3-14

Received: June 14, 2023
Revised: August 14, 2023
Accepted: August 29, 2023

Correspondence to:
Victoria L. Mango, MD, FSBI, Breast Imaging Service, MSK Ralph Lauren Center, Global Cancer Disparities Initiatives, Memorial Sloan Kettering Cancer Center, 300 East 66th Street, Suite 715, New York, NY 10065, USA
Tel. +1-646-888-4622
Fax. +1-646-888-4915
E-mail: mangov@mskcc.org

Introduction

Breast ultrasound (US) has several advantages over other imaging modalities, including its wide availability and lower cost. Additionally, US is a safe imaging modality as it neither involves ionizing radiation nor requires the administration of intravenous contrast agents, making it optimal for repeated examinations, particularly for young patients or pregnant women. It is commonly utilized as a supplemental screening tool, particularly for women with dense breasts in whom the sensitivity of mammography is decreased due to the masking effect of dense breast tissue [1,2]. Moreover, when a breast lesion is detected, US is often used in the diagnostic setting, further characterizing the lesion by describing the lesion's morphology on B-mode images including its' shape, margin, orientation, echo pattern and posterior acoustic features as well as its appearance while using different US techniques such as color, power Doppler, or elastography. US is also frequently used to guide minimally invasive procedures, such as needle biopsies of suspicious breast lesions or preoperative localization.

Nevertheless, US does have limitations, particularly handheld US (HHUS) which is highly operator dependent, contributing to great variation in the quality of US images depending on the skill and experience of the operator. Acquiring US images is time-consuming and interpretation requires

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2024 Korean Society of Ultrasound in Medicine (KSUM)



How to cite this article:
Brot HF, Mango VL. Artificial intelligence in breast ultrasound: application in clinical practice. Ultrasonography. 2024 Jan;43(1):3-14.

expertise and proficiency. US has faced criticism due to its relatively low specificity, which results in recalls and biopsies for benign lesions [3,4]. Recent research has demonstrated that the limitations of US in clinical practice may be improved through the use of artificial intelligence (AI) tools aimed at increasing the specificity of US and further improving its value as a screening and diagnostic modality. Additionally, with AI decision support for radiologists, AI applications may improve consistency in breast management recommendations by decreasing intraobserver and interobserver variability.

This review aims to highlight both current and emerging clinical applications of AI in handheld breast US.

Clinical AI Applications for Breast US

Breast Lesion Detection, Characterization, and Classification

Utilizing deep learning (DL) neural network techniques and instance segmentation, the application of AI research to clinical practice can potentially improve identification, characterization, and classification of US imaged breast lesions. When a breast lesion is detected, the suspicion for malignancy is primarily based on the radiologist's qualitative visual assessment guided by a standardized lexicon, part of the Breast Imaging Reporting and Data System (BI-RADS) Atlas [5], which contains a specifically developed section for breast US. Despite this, breast US has low specificity and low positive predictive values (PPVs). Large interobserver variability for lesion management recommendations has also been reported in clinical practice [6–11].

Intensively investigated in the field of breast imaging in recent years, several computer-aided detection/diagnosis (CAD) systems have emerged to enhance detection and diagnostic accuracy while reducing interpretation variability. Using machine learning and AI, these systems generate a probability of malignancy for a finding included in a user-selected region of interest (ROI), assisting the radiologist in the decision-making process when assigning a BI-RADS category and making management recommendations. As of today, there are local AI governance committees, with country-specific approved applications for clinical use. The United States Food and Drug Administration (FDA) approved a few AI-powered decision support applications for breast US clinical use (Table 1), which are commercially available on various vendor platforms [12].

BU-CAD (TaiHao Medical Inc.)

BU-CAD is an FDA-approved, commercially available AI software for computer assisted detection (CADe) and diagnosis (CADx DS) for breast US. To assist lesion detection, the CADe function identifies regions of interest (automated ROIs) of a suspicious lesion in up to two orthogonal US images. An adjunctive smart system

named breast free-hand US (BR-FHUS) which is comprised of two subsystems (BR-FHUS Navigation and Viewer) leverages the BU-CAD, allowing both real-time and batched recorded CADe assistance in identifying suspicious lesions for whole breast HHUS (Fig. 1A, B) [12]. Once a breast lesion is identified by either the human operator and/or the software and a ROI identified, the (CADx DS) component generates a numerical assessment of a lesion being malignant or benign based on its characteristics, termed by the manufacturer as "score of lesion characteristics (SLC)." The software also provides a corresponding BI-RADS category and descriptors including shape, orientation, margin, echo pattern, and posterior features (Fig. 2). The software was evaluated by Lai et al. [13] in a multi-reader study comparing the diagnostic performance and the interpretation time of breast US examinations between reading without and with the AI system as a concurrent reading aid. Statistically significant improvement of reader's diagnostic performance was achieved with the addition of BU-CAD CADx DS, increasing area under the receiver operating characteristic curve (AUC) from 0.758 to 0.829 ($P < 0.001$). Additionally, the readers' mean reading time has decreased from 30.15 seconds without the software to 18.11 seconds with the aid of the AI system, reflecting a significant decrease in interpretation time of nearly 40% ($P < 0.001$) [13]. The potential for AI to enable faster and more accurate interpretation by the reader could have tremendous clinical implications if these results are supported in future studies in a prospective clinical environment.

Koios Decision Support (DS) (Koios Medical Inc.)

The potential clinical impact of another breast US AI decision support system, Koios DS for breast, was recently evaluated by Mango et al. [14]. Koios DS is a software application designed to assist physicians in analyzing breast US images by generating a likelihood of malignancy for a user-selected ROI that contains a breast lesion (Fig. 3). In this multicenter retrospective study, 900 breast lesions seen on US images (470 benign and 430 malignant), were evaluated by 15 physicians with and without the software. The mean AUC with DS system alone (0.88; 95% confidence interval [CI], 0.86 to 0.91) and mean reader AUC with US plus DS (0.87; 95% CI, 0.84 to 0.90) were significantly higher than mean reader AUC with US only (0.83; 95% CI, 0.78 to 0.89; $P < 0.001$), demonstrating improved accuracy of sonographic breast lesion assessment using AI-based DS. Interobserver agreement, quantified by the Kendall τ -b correlation coefficient, was higher for US plus DS (0.68; 95% CI, 0.67 to 0.69) vs. US only (0.54; 95% CI, 0.53 to 0.55). Additionally, the integration of DS reduced the intra-observer variability, as evidenced by the decrease in the rates of cases with different BI-RADS assessments between the two reading sessions (10.8% vs. 13.6%, $P = 0.04$). There are inherent limitations of retrospective

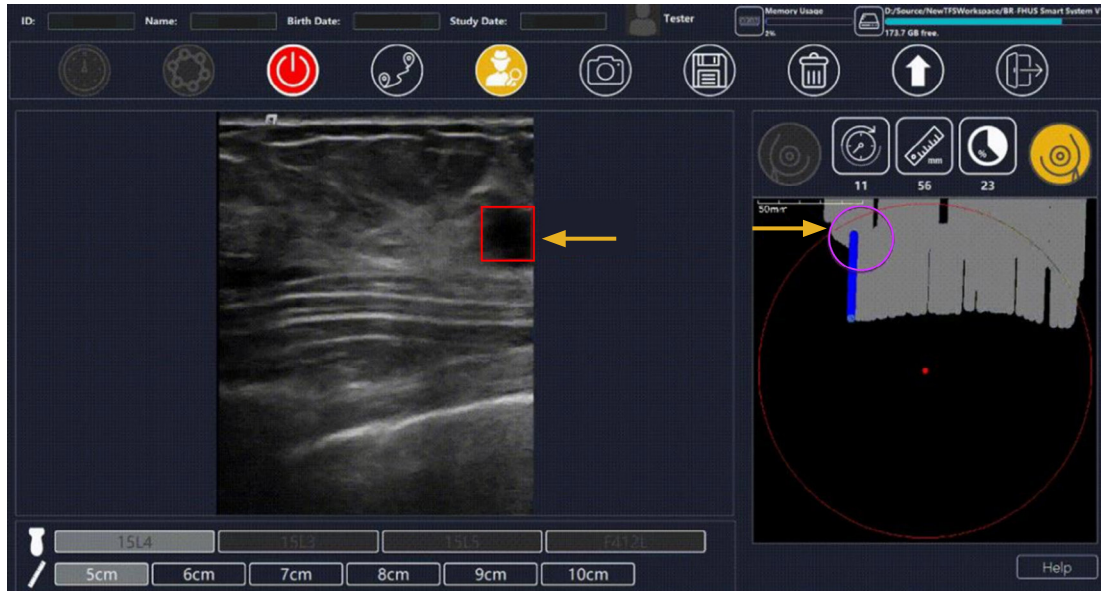
Table 1. Overview of commercially available artificial intelligence applications in breast handheld US

Study	Study design	No. of patients or lesions	AI technology	Clinical utility	Reported outcome
BU-CAD (TaiHao Medical Inc., Taipei, Taiwan)					
Lai et al. (2022) [13]	Retrospective, multi-reader and multi-case reader study	172 patients	Deep learning neural network techniques. Implements instance segmentation	- Detection CADE - assist users by generating automated ROIs of a single suspicious soft tissue lesion - Diagnosis CADx - generates a numerical assessment regarding likelihood of malignancy (score of lesion characteristics) and provides the correlating BI-RADS assessment	- Improves reader's diagnostic performance with significantly higher AUC and specificity - Significantly decreases average reading time
Koios DS (Koios Medical Inc.)					
Mango et al. (2020) [14]	Retrospective, multi-center, multi-reader	900 patients	Proprietary	Clinical DS - CADx software - generates a probability of malignancy for a breast lesion in a user selected ROI on static US images	- Improves accuracy of sonographic breast lesion assessment with significantly higher AUC - Reduces inter/intraobserver variability
Berg et al. (2021) [16]	Retrospective, multi-reader	319 lesions	Proprietary	Clinical DS - CADx software - generates a probability of malignancy for a breast lesion in a user selected ROI on static US images	- Original CADx did not substantially impact radiologists' interpretations - Improved performance and increased responsiveness of radiologists when CADx generated fewer false-positive cues
Browne et al. (2023) [17]	Retrospective, single institution	403 lesions	Proprietary	Software defined ROI of user identified lesions on static US images of biopsied lesions and analyzed the finding, generating the risk of malignancy using a similar scale to the BI-RADS classification	The use of AI decision support may contribute to the "triage" process assist radiologists to more confidently recommend a follow up US for a true BI-RADS 3 finding or upgrading it appropriately to a BI-RADS 4 with clear recommendation for biopsy
S-Detect (Samsung Healthcare, Seoul, South Korea)					
Zhao et al. (2022) [26]	Prospective, multicenter	757	DL based-CADx system constructed on convolutional neural network (CNN)	The software offers a dichotomous assessment, classifying sonographic screening-detected breast lesions observed on static US images as either "possibly benign" or "possibly malignant"	Significantly higher AUC and specificity, with no decrease in sensitivity
Wang et al. (2022) [30]	Meta-analysis of 11 studies	2,817 lesions	-	-	High diagnostic accuracy in distinguishing benign and malignant breast masses

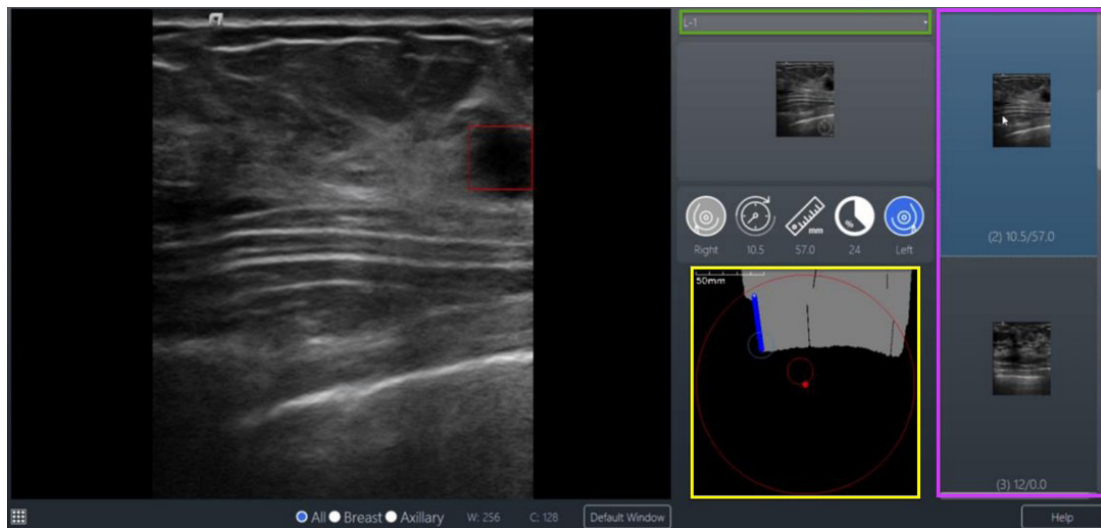
CADE, computer assisted detection; CADx, computer assisted diagnosis; AUC, area under the receiver operating characteristic curve; BI-RADS, Breast Imaging Reporting and Data System; DS, decision support; ROI, region of interest; US, ultrasound; AI, artificial intelligence; DL, deep learning.

breast US AI studies, as they do not replicate a true clinical environment in which lesions are often scanned in real time by the radiologist and evaluated in the context of patient symptoms, risk factors, and correlation with mammography, prior imaging, or both. Although AI-based DS reduced inter and intra-observer variability and improved correct assessment of sonographic breast lesions by most physicians, when the impact of DS on each reader's sensitivity

and specificity was analyzed, improvements appeared to depend on the reader's subspecialty. This finding was in line with results from a prior study by Chabi et al. [15], who evaluated the accuracy of CAD in breast US according to the radiologist's experience. The authors found improved sensitivity for junior radiologists but decreased specificity for experienced radiologists. This implies that although AI can enhance the diagnosis of malignant lesions when employed by



A



B

Fig. 1. BR-FHUS Navigation - an adjunct artificial intelligence tool to assist breast ultrasound screening.

A. The software provides a route map panel to display scanning information including probe location and scanning coverage (scanned areas of the breast indicated by gray bars). In addition, the system generates a "lesion detection module" to assist with real-time detection. Example of a suspicious lesion (red square) detected by the software with display of clock axis and distance to the nipple (purple circle) is shown. The images and cine loops are stored with their spatial position coordinates in DICOM format (figure provided by TaiHao Medical Inc. and used with permission). **B.** BR-FHUS Viewer assists physicians in reviewing a series of 2-D ultrasound images recorded by BR-FHUS Navigation. It supports the computer-aided detection method to detect breast lesions in the recorded images. An example of a snapshot taken by Navigation which is automatically showed in the Viewer so physicians can choose which lesion would appear in the report. By clicking a CAde button (not displayed on this image) the software automatically detects the suspicious areas which are displayed in the suspicious area list (purple rectangle), indicating breast laterality (green rectangle) and location within the breast (yellow rectangle) for each lesion. The user can capture the image which suits for reporting (in this case marked by blue background from the list area on the right side). Subsequently, the diagnostics results including DICOM images and reports are uploaded to PACS or stored in a local storage (figure provided by TaiHao Medical Inc. and used with permission). BR-FHUS, breast free-hand ultrasound; CAde, computer assisted detection; PACS, Picture Archiving and Communication System.

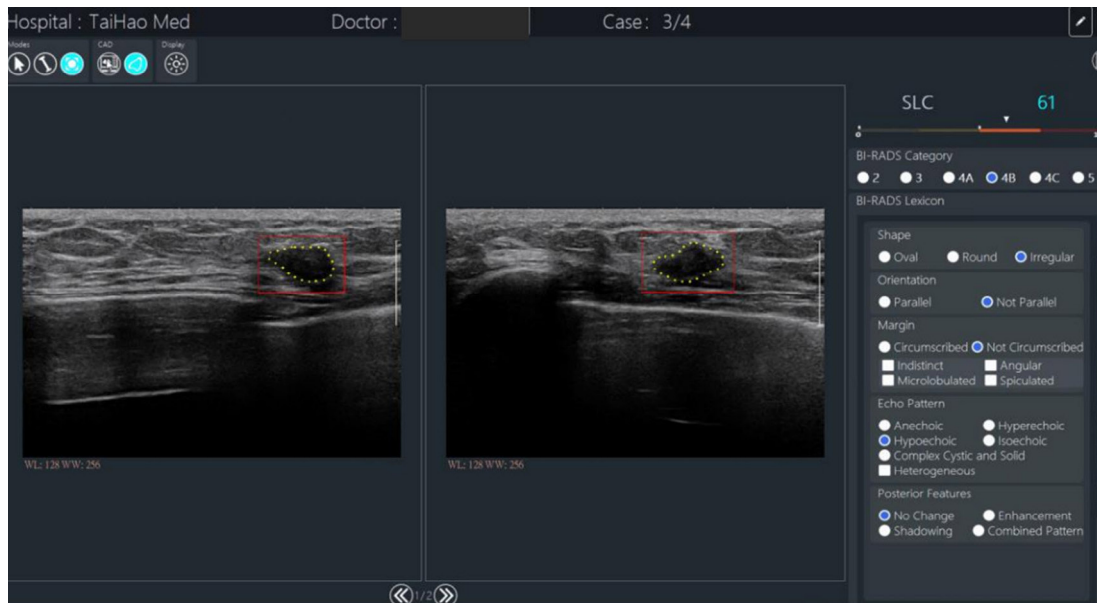


Fig. 2. Example of BU-CAD showing orthogonal B-mode ultrasound images of a breast mass (indicated by the red box), where the artificial intelligence decision support output (DS) displayed a high score of lesion characteristics (SLC=61), corresponding to a Breast Imaging Reporting and Data System (BI-RADS) assessment category 4B (figure provided by TaiHao Medical Inc. and used with permission).

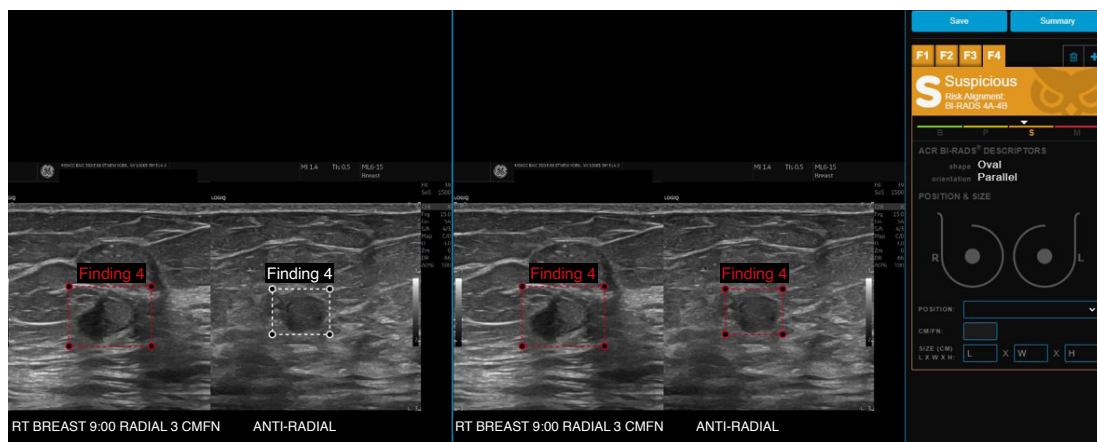


Fig. 3. Example of Koios decision support (DS) for Breast showing the B-mode ultrasound images of invasive carcinoma with micropapillary features in a 68-year-old patient. The artificial intelligence DS output is displayed in a graphical form on the right panel, with the DS-generated output (in this case correctly classified as "suspicious") and the confidence of assessment within that category as marked by the triangular marker.

junior radiologists, it might not surpass or enhance the precision of experienced, specialized breast radiologists [15]. This trend must be explored further and carefully considered when implementing breast US AI in clinical practice as the experience of the radiologists in a given practice may affect the impact of the AI system.

Additionally, Berg et al. [16] conducted a Koios reader study to assess the impact of original and artificially enhanced AI-based CADx on breast US interpretation. The study included 319 lesions identified on screening, with 88 (27.6%) being cancers. Nine

breast imaging radiologists reviewed orthogonal-paired US images, evaluating the findings both without and with CADx assistance in its original mode (AUC=0.77) and modified to high sensitivity (AUC=0.89) or specificity (AUC=0.86).

Results showed no overall improvement in accuracy when using the original CADx. However, when the DS outputs were adjusted to provide a binary categorization (benign or malignant) in high sensitivity or high specificity mode, all readers significantly enhanced their accuracies (average AUC increase of 0.045 for high

sensitivity mode and 0.071 for high specificity mode, $P < 0.001$). The authors concluded that radiologists acted more appropriately on CADx output when it presented fewer false-positive cues, perhaps indicating readers tend to trust a more specific tool. These results highlight that issues related to user trust of AI need to be considered in AI development, implementation, and radiologist training, as ultimately the human-AI interaction will affect the impact AI has on patient care.

The Koios software impact on the PPV of US-guided breast biopsies has also been investigated. Browne et al. [17] retrospectively compared pathology results of 403 biopsied breast lesions from a single institution against the original radiologist's BI-RADS classification based on breast US diagnostic images and the result of processing the same images through the AI algorithm. At the authors' institution, cytology is performed at the radiologist's discretion for some palpable breast lesions that are assessed as BI-RADS 3. While the BI-RADS Atlas indicates if biopsy is recommended a BI-RADS 4 or 5 assessment should be given, this study highlights the challenging clinical scenario of managing US findings that appear very likely to be benign but are palpable. The authors report that in their clinical practice if cytology renders indeterminate or suspicious results a core needle biopsy is recommended. According to study results, the software was more successful at determining which BI-RADS 3 lesions selected for biopsy by the radiologist actually required biopsy, potentially avoiding 44.1% (19/43) of biopsies yielding a benign result, without missing any cancers in this category. The use of AI DS may assist radiologists in this practice setting to more confidently recommend a 6-month follow-up US for a true BI-RADS 3 finding or upgrading it appropriately to a BI-RADS 4 with clear recommendation for biopsy. In this study, using Koios did not significantly increase the PPV for lesions categorized BI-RADS 4/5/6 by the reading radiologists and potentially would have missed 10 cancers, indicating that suspicious lesions categorized as BIRADS 4b and higher by reading radiologists, still warrants a biopsy if recommended by the human reader [17]. Research studies that focus only on lesions recommended for biopsy by the radiologist fail to capture cancers missed by the radiologist that may have been accurately categorized by the AI system. Such limitations should be considered when applying research results to real-world clinical practice.

S-Detect (Samsung Medison, Co., Ltd., Seoul, South Korea)

"S-Detect," is a commercially available CADx DS tool for breast US which employs a DL algorithm to offer a valuable second opinion, aiding operators in the interpretation and diagnosis of breast lesions. Once a clinician selects a breast lesion, S-Detect promptly generates an ROI encompassing the lesion, which the

operator can manually adjust as needed. The software analyzes the morphological characteristics of the lesion according to the BI-RADS lexicon, provides a detailed report of each US descriptor and combines the information with manual input from the operator regarding specific features (such as associated calcifications) and produces a dichotomized evaluation result of "possibly benign" or "possibly malignant" (Fig. 4) [18]. The impact of S-Detect on diagnostic performance has been investigated by multiple studies, reinforcing that CADx is a useful additional diagnostic tool in breast US for radiologists [19], with benefits varying depending on the radiologist's level of experience, primarily benefiting less experienced radiologists [20–25].

The S-Detect system was evaluated via a multicenter prospective study in China by Zhao et al. [26], investigating the feasibility of S-Detect in enhancing the diagnostic performance of breast US for screen-detected lesions. Analyzing 757 breast masses (460 benign, 297 malignant), S-Detect exhibited significantly higher AUC (0.83 [0.80–0.85]) and specificity (74.35% [70.10%–78.28%]) than radiologists reading conventional US (0.74 [0.70–0.77] and 54.13% [51.42%–60.29%] respectively; $P < 0.001$), with no decrease in sensitivity [26].

Improved specificity has important implications for patient care by reducing biopsies of benign lesions, which would spare the patient the associated discomfort, cost, anxiety, and time involved with undergoing a biopsy. Evaluating the potential of AI to reduce benign biopsies by downgrading a BI-RADS 4A findings to BI-RADS 3, Wang et al. [27] described two separate downgrading stratifications based on S-Detect AI assessments, according to whether the "possibly benign" AI-output was for both orthogonal US images vs. just one of the US images. While assessing 43 BI-RADS 4A lesions using the first strategy, the biopsy rate decreased from 100% to 67.4% ($P < 0.001$) with no missed cancers when both US images were assessed as "possibly benign"; however, when the assessment was "possibly benign" on just one of the images, the biopsy rate decrease was greater (from 100% to 37.2% [$P < 0.001$]) but two cancers were missed [27]. Comparable to the previously mentioned Koios readers study findings, intra-reader [28] and inter-reader agreement among radiologists regarding final BI-RADS assessment [20,29], and specific BI-RADS morphological descriptors, was improved by using S-Detect, with greater improvement noted among less experienced readers [29].

The overall diagnostic accuracy of S-Detect is best summarized by a meta-analysis by Wang and Meng [30], which included 11 studies in which 951 malignant and 1,866 breast masses were assessed. The results of the meta-analysis indicate that the system has high diagnostic accuracy in differentiating benign and malignant masses, with pooled sensitivity of 0.82 (95% CI, 0.74 to 0.88); pooled

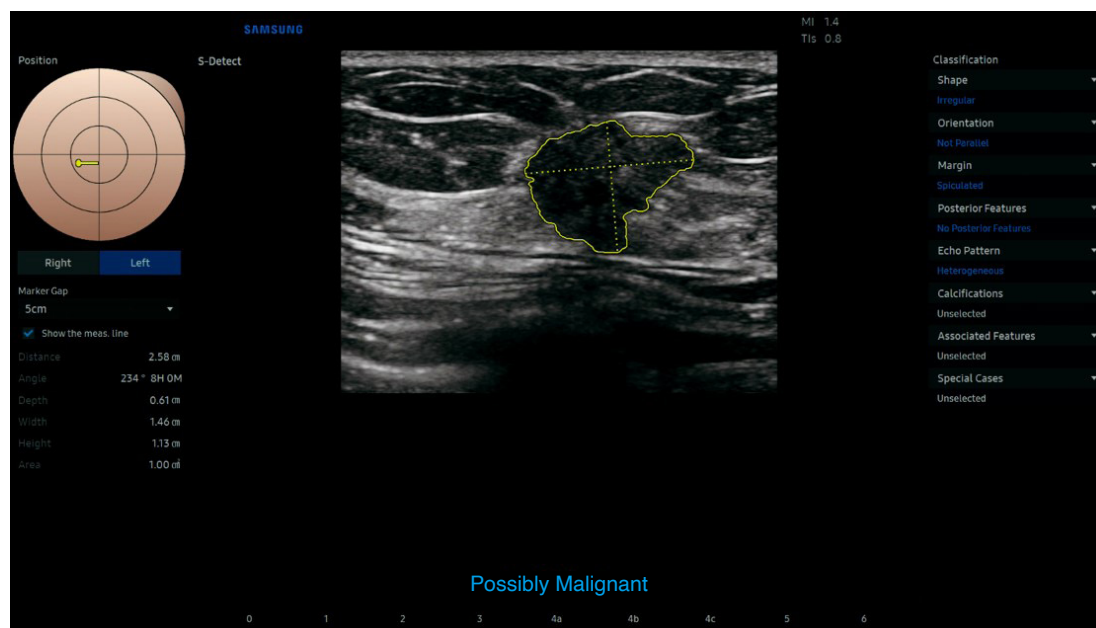


Fig. 4. Example of S-Detect showing the B-Mode ultrasound image of breast mass. The artificial intelligence decision support output outlined the region of interest around the mass margins (yellow line); classified descriptors as irregular shape, non-parallel orientation, spiculated margin, and heterogenous echo pattern. S-Detect produced a final assessment of "possibly malignant" supporting a recommendation to biopsy this finding (figure provided by Samsung Electronics Co., Ltd.).

specificity 0.83 (95% CI, 0.78 to 0.88) and AUC of 0.90 (standard error=0.0166) [30].

Prognostic AI Applications in Breast US

Prediction of Tumor Biology and Molecular Subtypes of Breast Cancer

Gene expression profiling has significantly influenced the comprehension of breast cancer biology and biomarkers. These encompass histologic grade, estrogen receptor, and progesterone receptor expressions, human epidermal growth factor receptor 2 (HER2) expression, and oncogene activity. These factors have been integrated into breast cancer staging by the American Joint Committee for Cancer (AJCC) [31]. Several studies have examined the prediction of breast cancer molecular subtypes based on AI systems for breast US.

The role of breast US AI in triple-negative cancer was investigated by Ma et al. [32] in a retrospective review of 600 patients with breast cancer, randomly divided into training (n=450) and testing (n=150) sets. Five AI models were trained based on clinical characteristics and imaging features of both mammography and US. One model excelled in distinguishing between triple-negative breast cancer and the other subtypes (AUC, 0.971; 95% CI, 0.947 to 0.995), significantly improving the accuracy, sensitivity, and specificity of four radiologists with the help of the model.

In a search of potential non-invasive preoperative methods to predict tumor molecular subtypes, a US-based assembled convolutional neural network multi-model was developed by Zhou et al. [33] trained on combined grayscale US images together with color Doppler and shear wave elastography features. The performance of the multi-model (macro-average AUC, 0.89–0.96) was superior to that of other models which were based on greyscale only (macro-average AUC, 0.73–0.75) or greyscale with color Doppler images (macro-average AUC, 0.81–0.84) in predicting four-classification breast cancer molecular subtypes. Surprisingly the authors reported that the model was also better than preoperative core needle biopsy (AUC, 0.89–0.99 vs. 0.67–0.82; $P < 0.05$), possibly due to partial sampling obtained by CNB which might not represent the entire lesion owing to the heterogeneity of breast cancer. The ability of AI to evaluate the entire lesion via the US images as opposed to partial sampling by core biopsy may prove to be a benefit and is an area of future research.

Prediction of Response to Neoadjuvant Chemotherapy in Breast Cancer

Neoadjuvant chemotherapy (NAC) is used to downstage locally advanced breast cancer and additionally serves as *in vivo* drug sensitivity testing. Evaluating the pattern of response to NAC can be used to further tailor systemic treatment [34]. Pathologic complete response (pCR) is associated with favorable disease-free and overall

survival and is the most commonly used endpoint in neoadjuvant trials [35]. While surgery is currently required to confirm a pCR post-NAC, with further study the omission of surgery could potentially be explored if pCR could be identified non-invasively. Although breast magnetic resonance imaging is currently the most accurate imaging modality for assessing pCR [36,37], the ability to predict pCR preoperatively using a less costly, less complex, and more accessible modality such as US has enormous clinical potential and economic benefits. This has encouraged the investigation of the added value of AI and deep learning radiomics (DLR) to US images in patients undergoing NAC. Jiang et al. [38] developed a novel preoperative pCR prediction model which is based on pre- and post-NAC US images, showing an AUC of 0.94 (95% CI, 0.91 to 0.77), outperforming two experienced radiologists.

In a prospective study by Gu et al. [39], two DLR models (DLR-2 and DLR-4) corresponding to two post-treatment time points (after second and fourth courses of chemotherapy) were used to assess response to NAC based on US images of 168 patients. DLR-2 achieved an AUC of 0.812 (95% CI, 0.77 to 0.85) and DLR-4 achieved an AUC of 0.937 (95% CI, 0.913 to 0.955). Furthermore, the deep learning radiomics pipeline model which is a stepwise prediction model that combines DLR-2 and DLR-4, successfully identified 90% (19/21 patients) of non-responders, triaging them to a different treatment strategy from which they could potentially benefit more.

Prediction of Axillary Nodal Metastases

The axillary lymph node status plays a crucial role in staging, prognosis, and treatment decisions for breast cancer patients [40]. Over time, axillary surgical management has evolved, moving away from routine axillary lymph node dissection (ALND) for all breast cancer patients. Instead, it focuses on avoiding ALND in cases of negative sentinel lymph node biopsy (SLNB) with similarly low axillary failure rates and significantly reduced lymphedema rates [41]. More recently the use of SLNB has expanded to women with low-volume nodal disease based on results of randomized controlled trials demonstrating equivalent locoregional control and survival [42,43]. Conducted by the American College of Surgeons Oncology Group (ACOSOG), the Z0011 trial showed that in women with early-stage, clinically node-negative (cN0) patients with less than three positive sentinel lymph nodes undergoing breast conserving surgery and whole breast radiation therapy, ALND can be omitted [44,45]. Large meta-analysis reported that preoperative axillary US combined with LN biopsy in the diagnostic workup of breast cancer patients will identify 50% of cases with axillary metastatic disease whereas in 25% of the patients with negative US-guided biopsy, a positive sentinel node may still be retrieved at SLNB [46]. The

current use of axillary imaging and image-guided biopsy in patients with breast cancer who meet Z0011 criteria or are undergoing NAC is institution- and surgeon-dependent, but also has the potential to overtreat the axilla and increase surgical morbidity [47]. Given the evolving surgical approach to axillary management, the role of axillary imaging must evolve in tandem, with the aim of improving diagnostic performance and supporting best patients' outcomes. Potentially, AI can assist in predicting axillary lymph nodes status based on sonographic images, enhancing the utility and accuracy of preoperative US.

A predictive model of lymph node metastasis in patients with breast cancer by using DL neural networks was developed and investigated by Zhou et al. [48]. Three representative deep convolutional neural networks (CNNs) models were evaluated to predict lymph node metastasis based on US images of primary breast cancer and of axillary lymph nodes. The performance of the models was compared with that of five radiologists while surgical pathologic results were used as the reference standard. Performance was analyzed in terms of accuracy, sensitivity, specificity, receiver operating characteristic curves, AUC, and heat maps. The best-performing CNN model was superior to a consensus of five radiologists, achieving higher sensitivity of 85% (95% CI, 70% to 94%) vs. 73% (95% CI, 57% to 85%; $P=0.17$) respectively and higher specificity of 73% (95% CI, 56% to 85%) compared to 63% (95% CI, 46% to 77%; $P=0.34$) achieved by the radiologists. Relatively high AUC of 0.89 (95% CI, 0.83 to 0.95) in the prediction of axillary lymph node metastasis was achieved by the model, demonstrating the feasibility of using CNNs to predict whether early primary breast cancer has metastasized to axillary lymph nodes.

Zheng et al. [49] investigated the incorporation of DLR to conventional US and shear wave elastography of breast cancer for predicting axillary lymph node status preoperatively in patients with early-stage breast cancer, reporting high diagnostic performance of combined clinical information and DLR output, with AUC of 0.90 (95% CI, 0.84 to 0.96) in comparison to lower AUC of 0.74 (95% CI, 0.69 to 0.78) achieved by routine axillary US done by radiologist. Nevertheless, there was no significant difference between DLR and radiologists' evaluation in terms of sensitivity, specificity, PPV or negative predictive value [49].

Breast US AI Challenges: Training and Implementation Considerations

AI, including its subfield of DL [50], holds the potential to revolutionize medical diagnostics. In an era of escalating imaging requirements and a scarcity of radiologists, the demand for automating the diagnostic process is on the rise. However, for the

successful utilization and integration of AI, it must match or surpass current technologies and healthcare professionals while also offering additional benefits such as swiftness, effectiveness, affordability, improved accessibility, and the preservation of ethical standards [51]. A significant limitation of DL in medical imaging is the requirement for substantial quantities of high-quality training data. This includes images with pixel-wise annotation and histological ground truth, or data with extended follow-up periods. Data augmentation results in the generation of more data, enhancing the model's capacity to handle diverse information in the testing set independently. Yet, the breakdown of this process could lead to overfitting, a notable machine learning hurdle which transpires when a model cannot extend its learned patterns beyond its training data [52]. The requirement for substantial high-quality datasets for the purpose of AI training within the field of breast imaging, may present a more prominent challenge in US, due to its relatively lower resolution compared to mammography [53]. Moreover, the generalizability of AI systems in breast US clinical practice may be limited by significant differences between image data used to train an algorithm, which may potentially be specific to the operator acquiring the US images, vendor or institution and those in real clinical practice [12].

AI - Breast US in Low Resource Settings

Increasing use of AI in radiology has raised concerns in high income countries that AI could potentially replace human radiologists, however in low and middle income countries (LMIC), where radiologists are limited or absent, AI has tremendous potential to bridge human resource gaps.

AI implementation in low resource areas of the world requires consideration of local stakeholder needs and available resources including technology infrastructure. RAD-AID International, a non-profit committed to enhancing medical imaging and radiology access in resource-limited global regions, has outlined clinical radiology education, infrastructure integration, and gradual AI implementation as pivotal components of a three-part strategy for LMIC AI adoption [54]. In a collaboration between RAD-AID International and Koios Medical (New York, NY), residents and local imaging professionals in low resource countries were introduced to the AI software and observed how AI outputs could be appraised for breast cancer biopsy recommendations and used as a DS tool [55,56].

Most importantly, AI may improve access to healthcare and bring radiology services to underserved areas with few or no radiologists. A recently prospective multicenter study in Mexico by Berg et al. [57] demonstrated that AI software applied to breast US images obtained with low-cost portable equipment and by minimally trained nonphysician research coordinators could accurately classify

and triage palpable breast masses in a low resource setting. In this study, targeted US was performed twice on women presenting with at least one palpable breast mass. First, orthogonal images with and without calipers were obtained of breast masses with use of the low-cost portable US, documenting any findings at the site of lump and adjacent tissue. The first 376 women were scanned by a specialist breast imaging radiologist and the subsequent 102 women were scanned by one of two nonphysician research coordinators who had been trained to use the portable US device by a validated 30-minute PowerPoint (Microsoft) presentation detailed by Love et al. [56]. Second, all women were also scanned with the use of standard-of-care (SOC) US, preformed and assessed by the specialized radiologist. Outputs of benign, probably benign, suspicious, and malignant were generated by an AI software (Koios DS version 3.x, Koios Medical). Seven hundred fifty-eight masses in 300 women were analyzed by the AI tool, of which 360 (47.5%) were palpable and 56 (7.4%) were malignant. The AI software correctly identified cancers in 47 or 48 of 49 women (96%–98%) with either portable US or SOC-US images, with AUCs of 0.91 and 0.95, respectively. Moderate specificity was achieved by the AI tool, correctly triaging 38% of the women with benign masses when analyzed masses imaged with portable low-cost US, and 67% of the women with benign masses imaged by a specialist radiologist using SOC Equipment ($P < 0.001$). As the authors acknowledge, while radiologists using low-cost portable HHUS could generate images of breast masses adequate for accurate AI classification, a suboptimal performance was achieved when AI applied to images obtained by minimally trained research coordinators using the same device, highlighting the need for greater training of personnel. It is also important to mention that Koios DS algorithms were not trained with images from the device used, and that training the software with images from low-cost portable US could potentially improve specificity.

The combination of AI with volume sweep imaging (VSI) US scans was assessed by Marini et al. [58] in a pilot study preformed to evaluate the possibility of inexpensive, fully automated breast US acquisition and preliminary interpretation without an experienced sonographer or radiologist. Expert-selected VSI images of exams obtained by medical students without prior experience, together with SOC images were input into an AI software (Samsung S-Detect for Breast) which output mass features and classification as "possibly benign" and "possibly malignant". An excellent diagnostic performance was obtained by the AI tool detecting malignant breast lesions with a sensitivity of 100% and specificity of 86%. Additionally, substantial agreement on the diagnosis of cancers, cysts, fibroadenomas, and lipomas was achieved between S-Detect interpretation of mass characteristics of VSI in relation to S-Detect

interpretation of SOC imaging (Cohen's $\kappa=0.79$; 95% CI, 0.65 to 0.94; $P<0.001$), expert VSI interpretation (Cohen's $\kappa=0.73$; 95% CI, 0.57 to 0.9; $P<0.001$), expert SOC-US interpretation (Cohen's $\kappa=0.73$; 95% CI, 0.57 to 0.9; $P<0.001$) and the pathological diagnosis (Cohen's $\kappa=0.8$; 95% CI, 0.64 to 0.95; $P<0.001$).

AI holds immense potential for enhancing healthcare delivery in resource-limited settings and addressing disparities. However, a careless deployment of AI might exacerbate radiology-related healthcare inequalities. By conscientiously adapting AI to personnel variations, disease prevalence, available radiology equipment, and by diligently addressing legal, regulatory, and ethical considerations [54], AI tools, both in the broader radiology domain and specifically in breast imaging, offer a pathway to better clinical education and improved imaging accessibility. This has the potential to enhance outcomes related to breast cancer in LMIC.

Conclusion

AI-based detection and diagnostic DS tools have the potential to serve an important clinical role in handheld breast US. There is evidence that breast US AI may soon be utilized in clinical practice for breast lesion detection, characterization, classification and to determine prognosis. Further prospective studies are necessary to comprehensively assess the influence of AI on actual clinical diagnostic performance and to develop effective strategies for integrating AI into real-world clinical settings, including examination of the tremendous potential of breast US AI in low resource settings.

ORCID: Hila Fruchtman Brot: <https://orcid.org/0000-0003-0533-5910>; Victoria L. Mango: <https://orcid.org/0000-0002-1844-7965>

Author Contributions

Conceptualization: Mango VL. Data acquisition: Brot HF, Mango VL. Data analysis or interpretation: Brot HF, Mango VL. Drafting of the manuscript: Brot HF. Critical revision of the manuscript: Mango VL. Approval of the final version of the manuscript: all authors.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

- Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007;356:227-236.
- Wanders JO, Holland K, Veldhuis WB, Mann RM, Pijnappel RM, Peeters PH, et al. Volumetric breast density affects performance of digital screening mammography. *Breast Cancer Res Treat* 2017;162:95-103.
- Lee JM, Arao RF, Sprague BL, Kerlikowske K, Lehman CD, Smith RA, et al. Performance of screening ultrasonography as an adjunct to screening mammography in women across the spectrum of breast cancer risk. *JAMA Intern Med* 2019;179:658-667.
- Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Bohm-Velez M, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA* 2008;299:2151-2163.
- Mendelson EB, Bohm-Velez M, Berg WA, Whitman GJ, Feldman MI, Madjar H, et al. ACR BI-RADS ultrasound. In: D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA, eds. *ACR BI-RADS Atlas, Breast Imaging Reporting and Data System*. Reston, VA: American College of Radiology, 2013;1-154.
- Berg WA, Bandos AI, Mendelson EB, Lehrer D, Jong RA, Pisano ED. Ultrasound as the primary screening test for breast cancer: analysis from ACIN 6666. *J Natl Cancer Inst* 2016;108:djv367.
- Bae MS, Han W, Koo HR, Cho N, Chang JM, Yi A, et al. Characteristics of breast cancers detected by ultrasound screening in women with negative mammograms. *Cancer Sci* 2011;102:1862-1867.
- 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.
- Berg WA, Blume JD, Cormack JB, Mendelson EB. Training the ACIN 6666 Investigators and effects of feedback on breast ultrasound interpretive performance and agreement in BI-RADS ultrasound feature analysis. *AJR Am J Roentgenol* 2012;199:224-235.
- Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. *Radiology* 2006;239:385-391.
- Abdullah N, Mesurolle B, El-Khoury M, Kao E. Breast imaging reporting and data system lexicon for US: interobserver agreement for assessment of breast masses. *Radiology* 2009;252:665-672.
- Villa-Camacho JC, Baikpour M, Chou SH. Artificial intelligence for breast US. *J Breast Imaging* 2023;5:11-20.
- Lai YC, Chen HH, Hsu JF, Hong YJ, Chiu TT, Chiou HJ. Evaluation of physician performance using a concurrent-read artificial intelligence system to support breast ultrasound interpretation. *Breast* 2022;65:124-135.
- Mango VL, Sun M, Wynn RT, Ha R. Should we ignore, follow, or biopsy? Impact of artificial intelligence decision support on breast ultrasound lesion assessment. *AJR Am J Roentgenol* 2020;214:1445-1452.
- Chabi ML, Borget I, Ardiles R, Aboud G, Boussouar S, Vilar V, et al. Evaluation of the accuracy of a computer-aided diagnosis (CAD) system in breast ultrasound according to the radiologist's experience. *Acad Radiol* 2012;19:311-319.
- Berg WA, Gur D, Bandos AI, Nair B, Gizienski TA, Tyma CS, et al.

- Impact of original and artificially improved artificial intelligence–based computer aided diagnosis on breast US interpretation. *J Breast Imaging* 2021;3:301-311.
17. Browne JL, Pascual MA, Perez J, Salazar S, Valero B, Rodriguez I, et al. AI: can it make a difference to the predictive value of ultrasound breast biopsy? *Diagnostics (Basel)* 2023;13:811.
 18. Zhang D, Jiang F, Yin R, Wu GG, Wei Q, Cui XW, et al. A review of the role of the S-Detect computer-aided diagnostic ultrasound system in the evaluation of benign and malignant breast and thyroid masses. *Med Sci Monit* 2021;27:e931957.
 19. Kim S, Choi Y, Kim E, Han BK, Yoon JH, Choi JS, et al. Deep learning-based computer-aided diagnosis in screening breast ultrasound to reduce false-positive diagnoses. *Sci Rep* 2021;11:395.
 20. Park HJ, Kim SM, La Yun B, Jang M, Kim B, Jang JY, et al. A computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of breast masses on ultrasound: added value for the inexperienced breast radiologist. *Medicine (Baltimore)* 2019;98:e14146.
 21. Nicosia L, Addante F, Bozzini AC, Latronico A, Montesano M, Meneghetti L, et al. Evaluation of computer-aided diagnosis in breast ultrasonography: improvement in diagnostic performance of inexperienced radiologists. *Clin Imaging* 2022;82:150-155.
 22. Zhao C, Xiao M, Jiang Y, Liu H, Wang M, Wang H, et al. Feasibility of computer-assisted diagnosis for breast ultrasound: the results of the diagnostic performance of S-detect from a single center in China. *Cancer Manag Res* 2019;11:921-930.
 23. Wei Q, Zeng SE, Wang LP, Yan YJ, Wang T, Xu JW, et al. The added value of a computer-aided diagnosis system in differential diagnosis of breast lesions by radiologists with different experience. *J Ultrasound Med* 2022;41:1355-1363.
 24. Di Segni M, de Soccio V, Cantisani V, Bonito G, Rubini A, Di Segni G, et al. Automated classification of focal breast lesions according to S-detect: validation and role as a clinical and teaching tool. *J Ultrasound* 2018;21:105-118.
 25. Cho E, Kim EK, Song MK, Yoon JH. Application of computer-aided diagnosis on breast ultrasonography: evaluation of diagnostic performances and agreement of radiologists according to different levels of experience. *J Ultrasound Med* 2018;37:209-216.
 26. Zhao C, Xiao M, Ma L, Ye X, Deng J, Cui L, et al. Enhancing performance of breast ultrasound in opportunistic screening women by a deep learning-based system: a multicenter prospective study. *Front Oncol* 2022;12:804632.
 27. Wang XY, Cui LG, Feng J, Chen W. Artificial intelligence for breast ultrasound: an adjunct tool to reduce excessive lesion biopsy. *Eur J Radiol* 2021;138:109624.
 28. Bartolotta TV, Orlando A, Cantisani V, Matranga D, lenzi R, Cirino A, et al. Focal breast lesion characterization according to the BI-RADS US lexicon: role of a computer-aided decision-making support. *Radiol Med* 2018;123:498-506.
 29. Bartolotta TV, Orlando AA, Di Vittorio ML, Amato F, Dimarco M, Matranga D, et al. S-Detect characterization of focal solid breast lesions: a prospective analysis of inter-reader agreement for US BI-RADS descriptors. *J Ultrasound* 2021;24:143-150.
 30. Wang X, Meng S. Diagnostic accuracy of S-Detect to breast cancer on ultrasonography: a meta-analysis (PRISMA). *Medicine (Baltimore)* 2022;101:e30359.
 31. Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ, et al. Breast cancer-major changes in the American Joint Committee on Cancer eighth edition cancer staging manual. *CA Cancer J Clin* 2017;67:290-303.
 32. Ma M, Liu R, Wen C, Xu W, Xu Z, Wang S, et al. Predicting the molecular subtype of breast cancer and identifying interpretable imaging features using machine learning algorithms. *Eur Radiol* 2022;32:1652-1662.
 33. Zhou BY, Wang LF, Yin HH, Wu TF, Ren TT, Peng C, et al. Decoding the molecular subtypes of breast cancer seen on multimodal ultrasound images using an assembled convolutional neural network model: a prospective and multicentre study. *EBioMedicine* 2021;74:103684.
 34. Heil J, Kuerer HM, Pfob A, Rauch G, Sinn HP, Golatta M, et al. Eliminating the breast cancer surgery paradigm after neoadjuvant systemic therapy: current evidence and future challenges. *Ann Oncol* 2020;31:61-71.
 35. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet* 2014;384:164-172.
 36. Berg WA, Gutierrez L, NessAiver MS, Carter WB, Bhargavan M, Lewis RS, et al. Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer. *Radiology* 2004;233:830-849.
 37. Yeh E, Slanetz P, Kopans DB, Rafferty E, Georgian-Smith D, Moy L, et al. Prospective comparison of mammography, sonography, and MRI in patients undergoing neoadjuvant chemotherapy for palpable breast cancer. *AJR Am J Roentgenol* 2005;184:868-877.
 38. Jiang M, Li CL, Luo XM, Chuan ZR, Lv WZ, Li X, et al. Ultrasound-based deep learning radiomics in the assessment of pathological complete response to neoadjuvant chemotherapy in locally advanced breast cancer. *Eur J Cancer* 2021;147:95-105.
 39. Gu J, Tong T, He C, Xu M, Yang X, Tian J, et al. Deep learning radiomics of ultrasonography can predict response to neoadjuvant chemotherapy in breast cancer at an early stage of treatment: a prospective study. *Eur Radiol* 2022;32:2099-2109.
 40. Winchester DP, Trabanino L, Lopez MJ. The evolution of surgery for breast cancer. *Surg Oncol Clin N Am* 2005;14:479-498.
 41. Veronesi U, Viale G, Paganelli G, Zurrada S, Luini A, Galimberti V, et al. Sentinel lymph node biopsy in breast cancer: ten-year results of a randomized controlled study. *Ann Surg* 2010;251:595-600.

42. Galimberti V, Cole BF, Zurrada S, Viale G, Luini A, Veronesi P, et al. Axillary dissection versus no axillary dissection in patients with sentinel-node micrometastases (IBCSG 23-01): a phase 3 randomised controlled trial. *Lancet Oncol* 2013;14:297-305.
43. Sola M, Alberro JA, Fraile M, Santesteban P, Ramos M, Fabregas R, et al. Complete axillary lymph node dissection versus clinical follow-up in breast cancer patients with sentinel node micrometastasis: final results from the multicenter clinical trial AATRM 048/13/2000. *Ann Surg Oncol* 2013;20:120-127.
44. Giuliano AE, McCall L, Beitsch P, Whitworth PW, Blumencranz P, Leitch AM, et al. Locoregional recurrence after sentinel lymph node dissection with or without axillary dissection in patients with sentinel lymph node metastases: the American College of Surgeons Oncology Group Z0011 randomized trial. *Ann Surg* 2010;252:426-432.
45. Giuliano AE, Ballman K, McCall L, Beitsch P, Whitworth PW, Blumencranz P, et al. Locoregional recurrence after sentinel lymph node dissection with or without axillary dissection in patients with sentinel lymph node metastases: long-term follow-up from the American College of Surgeons Oncology Group (Alliance) ACOSOG Z0011 randomized trial. *Ann Surg* 2016;264:413-420.
46. Diepstraten SC, Sever AR, Buckens CF, Veldhuis WB, van Dalen T, van den Bosch MA, et al. Value of preoperative ultrasound-guided axillary lymph node biopsy for preventing completion axillary lymph node dissection in breast cancer: a systematic review and meta-analysis. *Ann Surg Oncol* 2014;21:51-59.
47. Mango VL, Pilewskie M, Jochelson MS. To look or not to look? Axillary imaging: less may be more. *J Breast Imaging* 2021;3:666-671.
48. Zhou LQ, Wu XL, Huang SY, Wu GG, Ye HR, Wei Q, et al. Lymph node metastasis prediction from primary breast cancer US images using deep learning. *Radiology* 2020;294:19-28.
49. Zheng X, Yao Z, Huang Y, Yu Y, Wang Y, Liu Y, et al. Deep learning radiomics can predict axillary lymph node status in early-stage breast cancer. *Nat Commun* 2020;11:1236.
50. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436-444.
51. Aggarwal R, Sounderajah V, Martin G, Ting DS, Karthikesalingam A, King D, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ Digit Med* 2021;4:65.
52. Soffer S, Ben-Cohen A, Shimon O, Amitai MM, Greenspan H, Klang E. Convolutional neural networks for radiologic images: a radiologist's guide. *Radiology* 2019;290:590-606.
53. Becker AS, Mueller M, Stoffel E, Marcon M, Ghafoor S, Boss A. Classification of breast cancer in ultrasound imaging using a generic deep learning analysis software: a pilot study. *Br J Radiol* 2018;91:20170576.
54. Mollura DJ, Culp MP, Pollack E, Battino G, Scheel JR, Mango VL, et al. Artificial intelligence in low- and middle-income countries: innovating global health radiology. *Radiology* 2020;297:513-520.
55. Gonzalez Moreno IM, Trejo-Falcon J, Matsumoto MM, Huertas Moreno M, Martinez Galvez M, Farfan Quispe GR, et al. Radiology volunteers to support a breast cancer screening program in Peru: description of the project, preliminary results, and impressions. *Radiologia (Engl Ed)* 2022;64:256-265.
56. Love SM, Berg WA, Podilchuk C, Lopez Aldrete AL, Gaxiola Mascareno AP, Pathicherikollamparambil K, et al. Palpable breast lump triage by minimally trained operators in Mexico using computer-assisted diagnosis and low-cost ultrasound. *J Glob Oncol* 2018;4:1-9.
57. Berg WA, Lopez Aldrete AL, Jairaj A, Ledesma Parea JC, Garcia CY, McClennan RC, et al. Toward AI-supported US triage of women with palpable breast lumps in a low-resource setting. *Radiology* 2023;307:e223351.
58. Marini TJ, Castaneda B, Parker K, Baran TM, Romero S, Iyer R, et al. No sonographer, no radiologist: assessing accuracy of artificial intelligence on breast ultrasound volume sweep imaging scans. *PLOS Digit Health* 2022;1:e0000148.