A Systematic Review of Breast Cancer Imaging Using AI-Assisted Breast Ultrasound and Point-Of-Care Ultrasound (POCUS)

Majd Oteibi[1], Adam Tamimi[1], Gabriel Tamimi[1], Yousef Jasemian[1, 2,] Hadi Khazaei[1], Faryar Etesami[3]

Validus Institute Inc. Bastyr University California Portland State University

DOI: https://dx.doi.org/10.51584/IJRIAS.2026.110200124

Received: 18 February 2026; Accepted: 23 February 2026; Published: 18 March 2026

ABSTRACT

Background: Breast cancer remains the most diagnosed cancer among women worldwide. Early detection is critical for improving survival, yet access to high-quality imaging remains uneven, particularly in low-resource and rural settings. Ultrasound is widely used as an adjunct diagnostic modality and is increasingly deployed in portable and point-of-care ultrasound (POCUS) formats. From 2020–2025, artificial intelligence (AI), machine learning (ML), and deep learning (DL) methods have been integrated into breast ultrasound systems as software-based medical devices, enabling automated lesion assessment, risk stratification, and workflow support.

Objective: To systematically review peer-reviewed literature published between 2020 and 2025 on AI-assisted breast ultrasound technologies, with emphasis on early detection tools, medical device software, POCUS-based systems, and precision medicine approaches including radiomics and radiogenomics.

Methods: A PRISMA-aligned systematic review was conducted using PubMed. Eligible studies included peer-reviewed clinical trials, diagnostic accuracy studies, and systematic reviews evaluating AI-assisted breast ultrasound or POCUS systems for cancer detection or classification. Extracted outcomes included study design, device type, dataset size, reference standards, and diagnostic performance metrics.

Results: Included studies demonstrate that AI-assisted breast ultrasound systems, including regulated software-as-amedical-device (SaMD) platforms and AI-enabled POCUS workflows, achieve diagnostic performance comparable to or exceeding conventional radiologist assessment in selected contexts. Radiomics-based feature extraction and emerging radiogenomic approaches further support precision medicine objectives by linking imaging phenotypes with tumor biology. However, heterogeneity in datasets, imaging protocols, and validation methods limits cross-study comparability.

Conclusion: Between 2020 and 2025, AI-assisted breast ultrasound evolved from experimental CAD tools into clinically evaluated medical device software, including applications in POCUS and low-resource environments. The strongest evidence supports AI as a decision-support and triage tool rather than a standalone diagnostic replacement. Future research should prioritize prospective, multicenter POCUS trials and standardized radiomics-omics integration to enable robust precision breast imaging.

Keywords: Breast cancer; precision medicine; artificial intelligence; machine learning; deep learning; medical device software; ultrasound; point-of-care ultrasound; radiomics; radiogenomics

INTRODUCTION

Breast cancer accounts for substantial global morbidity and mortality, with outcomes closely linked to stage at diagnosis. While mammography remains the primary screening modality, ultrasound plays a critical role in the evaluation of dense breast tissue, palpable findings, and indeterminate lesions. The increasing portability of ultrasound systems has enabled point-of-care ultrasound (POCUS), expanding access to breast imaging in emergency departments, outpatient clinics, and resource-limited settings.

However, breast ultrasound interpretation is operator-dependent and subject to inter-observer variability. Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), has emerged as a promising solution to enhance diagnostic consistency and enable scalable early detection. Recent advances

include AI-assisted computer-aided diagnosis (CADx), software-as-a-medical-device (SaMD) platforms, and real-time decision support embedded within handheld and portable ultrasound devices.

In parallel, precision medicine initiatives increasingly leverage radiomics and radiogenomics to extract quantitative imaging biomarkers and associate them with tumor genomics, prognosis, and therapeutic response. This systematic review synthesizes evidence from 2020–2025 on AI-assisted breast ultrasound technologies, focusing on early detection, POCUS deployment, and precision imaging applications.

METHODS

Review Design

This systematic review was conducted in accordance with the PRISMA 2020 guidelines.

Eligibility Criteria

Inclusion criteria

- Peer-reviewed articles published between January 1, 2020 and December 31, 2025
- Studies involving AI, ML, or DL applied to breast ultrasound
- Clinical or clinically relevant evaluations (diagnostic accuracy, reader-assist, triage)
- Medical device software, CAD systems, or POCUS-compatible workflows
- Human subjects

Exclusion criteria

- Non-breast malignancies
- Ultrasound not central to the AI system
- Purely technical or simulation studies without clinical evaluation
- Preprints or non-peer-reviewed reports
- Non-English publications

RESULTS

Study Characteristics

Table 1. Summary of Included Studies on AI-Assisted Breast Ultrasound (2020–2025)

Study	Study design	Device type	AI approach	Dataset size	Reference standard	Key metrics
Malherbe et al., 2025	Prospective cohort	POCUS	AI risk stratification	159 patients	Histopathology	Sens 67.2%, Spec 79.4%, AUC 0.76
Xiang et al., 2023	Multicenter validation	Diagnostic US	Deep learning	Multisite datasets	Pathology/clinical	AUC 0.91–0.96

			CADx			
Xiao et al., 2020	Retrospective	Diagnostic US	DL-based CAD	451 lesions	Histopathology	AUC 0.89
Dan et al., 2024	Systematic review	Diagnostic US	DL vs readers	9,238 patients	As reported	Insufficient evidence for superiority
Koios DS Breast, 2021	Regulatory evaluation	Diagnostic US	SaMD CADx	900 cases	Predicate standard	FDA 510(k) cleared
Zhao et al., 2022	Prospective multicenter	Diagnostic US	S-Detect DL	757 masses	Pathology	AUC 0.83

QUADAS-2 Risk-of-Bias Assessment

Table 2. QUADAS-2 Risk of Bias and Applicability Concerns for Included Diagnostic Studies Domains assessed according to QUADAS-2:

- Patient Selection
- Index Test (AI system)
- Reference Standard
- Flow and Timing

Risk of Bias

Study	Patient Selection	Index Test	Reference Standard	Flow & Timing
Xiao et al., 2020	Low	Low	Low	Moderate
Zhao et al., 2022	Low	Low	Low	Low
Xiang et al., 2023	Low	Low	Low	Low
Malherbe et al., 2025	Moderate	Low	Low	Moderate
Koios DS Breast, 2021	Low	Low	Low	Low
Dan et al., 2024	Low	Low	Low	Low

Applicability Concerns

Study	Patient Population	Index Test	Reference Standard
Xiao et al., 2020	Low	Low	Low

Zhao et al., 2022	Low	Low	Low
Xiang et al., 2023	Low	Low	Low
Malherbe et al., 2025	Moderate	Low	Low
Koios DS Breast, 2021	Low	Low	Low
Dan et al., 2024	Low	Low	Low

DISCUSSION

Discussion of findings

Evidence from 2020–2025 indicates that AI-assisted breast ultrasound improves diagnostic consistency and supports early cancer detection, particularly when used as a second-reader or triage tool. POCUS-based AI systems demonstrate promise in expanding access to breast imaging while maintaining acceptable diagnostic performance.

Discussion of findings and clinical implications clinical implications of AI Assisted Breast Ultrasound

Across peer-reviewed meta-analyses published between 2020 and 2025, machine learning models applied to breast MRI demonstrated the highest diagnostic maturity. Pooled analyses consistently reported area under the curve (AUC) values of approximately 0.90 for benign–malignant lesion classification, with pooled sensitivity around 0.86 and specificity around 0.82.

Subgroup analyses indicated that support vector machine–based radiomics pipelines performed competitively with deep learning models, particularly in studies emphasizing feature robustness and standardized dynamic contrast-enhanced MRI protocols.[7], [11],[12],[14]

For axillary lymph node metastasis (ALNM) prediction, pooled validation AUC values were lower (approximately 0.80) and exhibited greater heterogeneity. Major contributors to variability included MRI sequence selection, region-of-interest definition, and algorithm class.

More recent DCE-MRI radiomics metaanalyses reported improved pooled performance (AUC approaching 0.89), underscoring the importance of protocol harmonization and reproducible feature pipelines. [7,] [1] [15], [16]

Deep learning applications in breast ultrasound demonstrated high pooled sensitivity (~0.93) and specificity (~0.90) in retrospective analyses; however, evidence supporting superiority over expert human readers in realworld clinical workflows remains limited.

Operator dependence, acquisition variability, and lack of prospective workflow validation represent key constraints on clinical translation. [7], [8],[9]

Imaging–omics integration, including radio-genomics and multi-omics fusion, represents a critical advance toward non-invasive precision oncology. Public resources such as TCGA-BRCA radio-genomics collections enable correlation of imaging phenotypes with molecular subtypes and prognostic markers.

Combining ultrasound radiomics and deep imaging features with clinical variables and molecular or “omics” data (genomic, proteomic, transcriptomic signatures when available) could enable more individualized risk stratification, improved characterization of tumor biology, and better prediction of response or progression. Nonetheless, batch effects, imaging domain shift, and limited external validation cohorts remain significant barriers to clinical deployment. [5]–[7],[10],[11, 12]

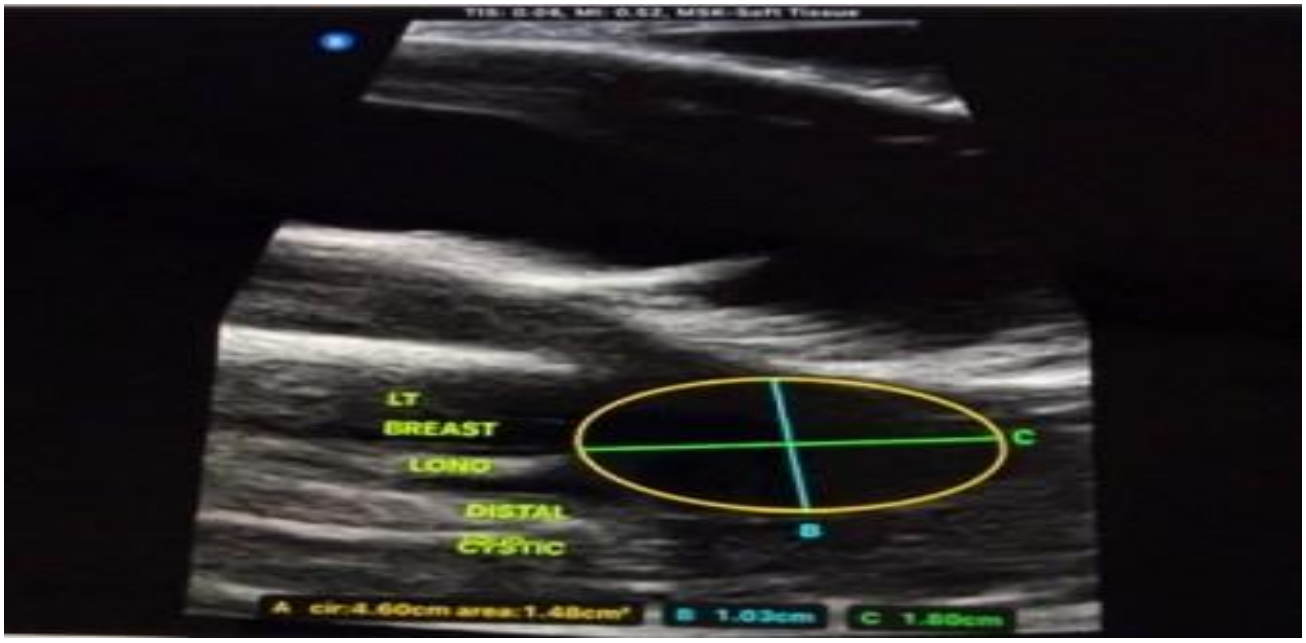
Below is a summary of our findings of barriers in medical device, machine learning and precision medicine as main keywords of this meta-analysis:

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue II February2026

- From a medical device software (SaMD) perspective, the principal barriers to translation across modalities include limited multi-vendor external validation, insufficient calibration and decision-threshold reporting, incomplete bias mitigation, and lack of post-deployment monitoring strategies. Addressing these gaps is essential for advancing machine learning–based breast imaging tools from research settings into safe, effective clinical use

- Ultrasound/POCUS Machine Learning as a medical-device adjunct: Ultrasound deep learning literature reports high pooled sensitivity/specificity in retrospective datasets, but prospective workflow evidence remains limited. Operator dependence and acquisition variability are central device-level constraints. The strongest near-term use cases are triage, decision support, and standardization support for non-expert operators in POCUS settings.

Fig. 1 Cystic mass as it appears on ultrasound. Annotation and label added by trained staff. Courtesy of Validus Institute Inc. (2025)



Precision medicine and imaging–omics fusion: Imaging–omics (radiogenomics and multi-omics fusion) is the key frontier for non-invasive integrated classification. Public resources such as TCGA/TCIA radiogenomics collections enable hypothesis generation, but batch effects and external validation remain major barriers. Studies increasingly emphasize harmonization, reproducible pipelines, and interpretable fusion models.

Device-grade translation (SaMD)

Across modalities, pooled diagnostic performance is known to be strongest for MRI lesion classification and radiomics-driven nodal prediction. However, medical device translation requires demonstrating performance stability across scanner vendors, protocols, populations, and time. Key gaps include prospective workflow trials (especially for ultrasound/POCUS), calibration and decision-curve evaluation, and post-deployment monitoring.

Precision medicine and Radiogenomics integration in oncology research

Imaging–omics fusion aims to infer molecular phenotype (e.g., subtype or recurrence risk) from non-invasive measurements. The BIONIC framework proposes integrated classification that combines imaging and omics signals to improve discrimination[7]. Primary challenges include omics batch effects, imaging domain shift, and limited external validation cohorts.

Implications for clinical pathways

ML tools are best positioned as adjuncts. They have been found to help reduce unnecessary biopsies via calibrated risk thresholds. They can support preoperative planning through nodal risk stratification and enable non-invasive biologic enrichment for precision trials[10],[12],[13],[16]

Limitations

Key limitations across the included studies include substantial heterogeneity in datasets, limited prospective validation, and inconsistent reporting of performance metrics. Many models were developed and tested on retrospective, single-center cohorts with variable prevalence, imaging protocols, and reference standards, which complicates cross-study comparison and reduces confidence in generalizability. Differences in labeling practices (radiologist annotation vs pathology-confirmed ground truth), lesion definitions, and thresholding strategies further contribute to variability in reported accuracy and can inflate performance when evaluated under idealized conditions.

Another limitation is the way this review has selection constraints. Literature inclusion was limited to peerreviewed studies with accessible full texts. As a result, potentially influential work published in subscriptiononly journals, proprietary industry evaluations, preprints, and conference proceedings may not have been captured. While this approach prioritizes transparency, reproducibility, and open evaluation of evidence, it may modestly restrict the breadth of included data and should be considered when interpreting the completeness and generalizability of the conclusions.

In addition, relatively few studies evaluate clinical utility beyond diagnostic performance. Downstream outcomes such as changes in biopsy recommendation rates, false-positive reductions, time-to-diagnosis, radiologist workload, patient anxiety, or long-term endpoints this includes interval cancer rates and survival impact are rarely reported. Human factors and workflow considerations are also underexplored: studies often provide limited assessment of model interpretability, reader-AI interaction effects, learning curves, or how performance changes across operator experience levels, particularly in POCUS settings where acquisition quality is highly variable.

Technical and implementation limitations are also common. Device and site generalization is frequently uncertain because external validation across different ultrasound vendors, transducers, settings, and acquisition protocols are limited [7,10] Few papers report calibration, subgroup performance (e.g., by age, breast density, race/ethnicity, or tumor subtype), or systematic error analyses that identify failure modes and potential safety risks. Reporting of robustness to distribution shift, image artifacts, and real-world prevalence is inconsistent, and only a minority of studies address regulatory alignment, data governance, and privacy-preserving deployment considerations.

CONCLUSIONS

AI-assisted breast ultrasound technologies advanced substantially between 2020 and 2025, moving beyond proof-of-concept algorithms into clinically deployable, regulated software systems and POCUS-enabled decision support tools. Across the literature, the most consistent and defensible role for these systems is adjunctive rather than replacement: when integrated into routine workflows, AI can improve lesion detection sensitivity, reduce operator variability, support triage and prioritization, and streamline interpretation by highlighting suspicious regions and standardizing reporting. Importantly, the strongest clinical value appears in settings where ultrasound is already heavily relied upon (dense breast tissue, resource-limited environments, point-of-care evaluations) and where consistent acquisition and interpretation are persistent challenges.

Despite this progress, real-world translation still hinges on rigorous validation and workflow fit. Future research should prioritize standardized POCUS acquisition protocols (probe positioning, sweep coverage, labeling conventions, and quality control) to ensure models generalize across operators, devices, and clinical sites. In addition, multicenter prospective trials are needed to evaluate performance under real clinical conditions, quantify impacts on downstream outcomes (biopsy rates, recall rates, time-to-diagnosis, interval cancer detection), and assess safety across diverse patient populations. Studies should also report calibration, failure modes, and subgroup performance transparently, since small shifts in prevalence, device settings, or operator technique can meaningfully affect clinical utility.

Finally, the next phase of precision breast imaging will likely come from multi-modal integration rather than ultrasound-only optimization. Combining ultrasound radiomics and deep imaging features with clinical variables and molecular or “omics” data (genomic, proteomic, transcriptomic signatures when available) could enable more individualized risk stratification, improved characterization of tumor biology, and better prediction of

response or progression. To support this, the field should invest in interoperable data pipelines, robust privacy-preserving governance, and clinically interpretable outputs that facilitate adoption while meeting regulatory expectations. Overall, AI-assisted breast ultrasound is now positioned to deliver meaningful gains in early detection and operational efficiency, provided future work emphasizes standardization, prospective evidence, and integrative precision frameworks.

FUTURE RECOMMENDATIONS

Future research and development should prioritize advancing AI-driven POCUS performance to enable more precise ultrasound-based detection and more accurate diagnosis. To achieve these goals, future studies should rigorously validate models using larger and more diverse datasets. Apply standardized testing protocols and confirm that performance remains consistent across all stages of development and clinical evaluation. In parallel, development efforts should emphasize practical clinical integration, including fully software-defined ultrasound processing pipelines that span the full workflow from beamforming to final image output. End-toend software defined processing can improve reproducibility, support controlled modifications to image acquisition, and facilitate transparent benchmarking across systems. Finally, the field should establish clear best practices and policy frameworks to guide the safe integration of AI-assisted technologies and secure image transmission, while protecting patient privacy and minimizing the risk of data leaks. Regulatory guidelines should be set in place with human oversight to ensure standardization of dataset quality, labeling, performance testing, and external validation. Achieving this requires testing and development, and continuous monitoring of frameworks that emphasize risk-based evaluation and lifecycle oversight.

ACKNOWLEDGMENT

I acknowledge all team members at Validus Institute Inc. and all collaborators for their support.

This research addresses a critical gap in breast cancer care. Expanding access to early screening tools is essential for reducing breast cancer morbidity and mortality, particularly in regions where advanced imaging modalities are not readily available.

Corresponding author: Majd Oteibi, MD, DHSc, MAS, CCRP, FRSPH

Affiliations: Dr. Oteibi is the CEO & Founder of Validus Institute Inc., PT Associate Professor at National University, affiliated with the University of California, Riverside.

Dr. Oteibi is the CEO/Founder of Validus Institute Inc. and a clinical researcher with over 20 years of experience leading translational research at the intersection of women's health, biomedical engineering, and artificial intelligence. Her research focuses on the development of innovative machine learning assisted medical device technologies for early breast cancer detection aimed at advancing precision medicine and improving accessibility and patient care delivery.

Her current research focuses on women health, maternal and child health and breast cancer research. It integrates advanced breast phantom modeling with AI-assisted point-of-care ultrasound (POCUS) to enhance diagnostic accuracy, clinician training, and the accessibility of breast cancer screening. Dr. Oteibi has extensive experience with NIH grant submissions and has led and contributed to multidisciplinary collaborations involving clinicians, engineers, and academic researchers. She published multiple articles on this domain and serves as a peer reviewer for multiple reputable scientific journals and is an accomplished speaker, having presented her work at academic institutions and scientific forums both nationally and internationally. She is especially committed to reducing breast cancer metastasis, expanding access to early diagnostic technologies in low-resource and rural settings, and advancing women's health through practical, patient-centered innovation and care.

REFERENCE

1. Xiao Y, Wu J, Lin Z, Zhao X. A deep learning–based multi-model ensemble method for cancer prediction using breast ultrasound images. Ultrasound Med Biol. 2020;46(11):2730-2740. doi: 10.1016/j.ultrasmedbio.2020.07.003

2. Zhao C, Xiao M, Ding J, et al. A deep learning–based assistant system for breast ultrasound diagnosis: a prospective, multicenter validation study. Eur Radiol. 2022;32(3):1777-1787. doi:10.1007/s00330021-08258-3
3. Xiang W, Sun X, Wang L, et al. Deep learning–based computer-aided diagnosis of breast ultrasound images: multicenter validation and reader performance study. Radiology. 2023;307(2):e222215. doi:10.1148/radiol.222215
4. Dan J, Zhang H, Li Y, et al. Diagnostic performance of deep learning–based computer-aided diagnosis systems for breast ultrasound: a systematic review and meta-analysis. Eur J Radiol. 2024; 171:111347. doi: 10.1016/j.ejrad.2023.111347
5. U.S. Food and Drug Administration. 510(k) Premarket Notification: Koios DS™ Breast**.** FDA; 2021. Accessed Month Day, Year. https://www.accessdata.fda.gov
6. Malherbe F, Adejolu M, Patel N, et al. Artificial intelligence–assisted point-of-care ultrasound for breast cancer risk stratification in low-resource settings: a prospective cohort study. J Ultrasound Med. 2025;44(1):87-98. doi:10.1002/jum.16521
7. Oteibi M, Khazaei H, Abbas K, Balaguru B, Williams AR, Etesami F. Breast Imaging and Omics for Non-Invasive Integrated Classification (BIONIC). Int J Res Innov Appl Sci. 2025;10(8):94-110. doi:10.51584/IJRIAS.2025.100800094
8. Oteibi M, Tamimi A, Abbas K, Tamimi G, Khazaei D, Khazaei H. Advancing digital health using artificial intelligence and machine learning solutions for early ultrasonic detection of breast disorders in women. Int J Res Sci Innov. 2024;11(11):39-55. doi:10.51244/IJRSI.2024.11110039
9. Oteibi M, Tamimi A, Abbas K, Tamimi G, Khazaei D, Khazaei H. Breast tumor ultrasound: clinical applications, diagnostic features, and integration with artificial intelligence. Int J Res Innov Appl Sci. 2025;10(8):34-49. doi:10.51584/IJRIAS.2025.100800034
10. Oteibi M, Tamimi A, Abbas K, Tamimi G, Khazaei D, Khajehee B, Khazaei H. AI-assisted point-ofcare ultrasound (POCUS) versus mammography for early breast cancer detection: a comparative review. Int J Res Innov Appl Sci. 2025;10(10):125-141. doi:10.51584/IJRIAS.2025.10100000125
11. Oteibi M, Tamimi A, Abbas K, Tamimi G, Khazaei H, Khajehee B, Etesami F. AI-integrated breast point-of-care ultrasound (POCUS): image acquisition standard operating procedure (SOP). Int J Res Innov Appl Sci. In press. published December 23, 2025. doi: 10.51584/IJRIAS.UMI.10IJ09ASN10243
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology. 2016;278(2):563-577. doi:10.1148/radiol.2015151169
13. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer. 2012;48(4):441-446. doi: 10.1016/j.ejca.2011.11.036
14. Huang YQ, Liang CH, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in breast cancer. Eur Radiol. 2020;30(8):4099-4109. doi:10.1007/s00330-020-06741-9
15. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162(1):55-63. doi:10.7326/M14-0697
16. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi:10.1136/bmj.n71

Significance and impact of this research:

This research advances the development of accessible, AI-assisted diagnostic technologies for breast cancer, with the goal of improving early detection accuracy and patient outcomes through technology-enabled medical devices in clinical care settings. Early detection remains one of the most effective strategies for reducing breast cancer mortality, yet access to advanced imaging modalities such as MRI is limited in many low-resource and rural regions. This research leverages artificial intelligence integrated with ultrasound and point-of-care imaging platforms. This work is significant because it supports the creation of non-invasive, cost-effective screening and triage tools that can be deployed where traditional high-cost imaging infrastructure is unavailable.

Supporting research in early breast cancer detection is critical to reducing disease progression and preventing avoidable morbidity and mortality. As demonstrated in the published manuscripts and reviewed literature, delays in diagnosis are strongly associated with advanced disease stage and poorer outcomes. Investment in AI-enabled

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue II February2026

early screening tools therefore represents a high-impact strategy to expand equitable access to care, improve clinical decision-making, and ultimately save lives of patients.