REVIEW ARTICLE
https://doi.org/10.29289/2594539420260013

The effectiveness of artificial intelligence to support the diagnosis of breast lesions by ultrasound: an integrative review

Mariana Loureiro Lemos[1], Bianca Gutfilen[2], Maria Julia Gregorio Calas[3*]

ABSTRACT

Breast cancer remains the most common malignant neoplasm among women worldwide, and early detection and diagnostic accuracy are fundamental to reducing mortality. Although mammography is the standard screening method, breast ultrasound plays an essential complementary role, despite being operator-dependent, with high inter-observer variability and lower specificity. In this context, artificial intelligence emerges as a promising tool for improving image interpretation, serving as a clinical decision-making aid. This integrative review analyzed comparative studies obtained between January 2016 and January 2024 from the PubMed/MEDLINE, LILACS, SciELO, and Connected Papers databases, including original articles with histopathological confirmation as the gold standard and excluding other imaging methods and studies outside the scope. A total of 14 studies (2016–2023) were included, involving 3,144 patients and 3,524 lesions (2,318 benign and 1,206 malignant), evaluated by approved artificial intelligence systems (Koios DS and S-Detect), with heterogeneity regarding evaluator experience and methodological criteria. The results demonstrate that artificial intelligence exhibits high diagnostic performance, with sensitivity and area under the curve comparable to those of experienced radiologists, in addition to improving standardization, reducing inter-observer variability, and increasing accuracy—especially among less experienced professionals—with high negative predictive value and the potential to reduce unnecessary biopsies. However, specificity is variable, and when artificial intelligence is used alone, it may be lower than that of specialists. The combined use of a radiologist and artificial intelligence yielded better results. Methodological and technical limitations—including a predominance of retrospective studies, data heterogeneity, and system constraints in evaluating certain findings—hinder the generalizability of results. Artificial intelligence emerges as a promising complementary tool in breast ultrasound, with the potential to improve diagnostic accuracy and clinical management. Prospective, multicenter, and standardized studies are needed to consolidate its application in clinical practice.

KEYWORDS: artificial intelligence; breast; biopsy; breast cancer.

INTRODUCTION

Breast cancer remains the most common malignant neoplasm among women and one of the leading causes of cancer mortality worldwide. Early detection and diagnostic accuracy are directly associated with reduced mortality, rendering the continuous improvement of clinical imaging techniques imperative[1–3].

Although mammography is recognized as the standard method for population-based breast cancer screening and has a proven impact on reducing mortality, breast ultrasound plays a key role as a complementary method in the diagnosis of breast lesions. Breast ultrasound offers significant advantages, such as widespread availability, lower cost, the absence of ionizing radiation, no breast compression, and no need for intravenous contrast, making it particularly suitable for repeated evaluations in young patients, pregnant women, and women with dense breasts. Furthermore, it is the method of choice for evaluating axillary nodes, further characterizing findings detected on mammography or magnetic resonance imaging, clarifying palpable lesions, differentiating solid from cystic lesions, and guiding minimally invasive procedures[1–4].

[1]Universidade Federal do Rio de Janeiro, Master's Degree Program – Rio de Janeiro (RJ), Brazil.
[2]Universidade Federal do Rio de Janeiro, Doctorate Degree Program – Rio de Janeiro (RJ), Brazil.
[3]Universidade Estácio de Sá, Instituto de Educação Médica – Rio de Janeiro (RJ), Brazil.
***Corresponding author:** mariajulia.calas@gmail.com
Editor: Mariana Loureiro Lemos
Received on: 04/27/2026. **Accepted on:** 04/08/2026. **Published on:** 11/11/2026.
How to cite: Lemos ML, Gutfilen B, Calas MJG. The effectiveness of artificial intelligence to support the diagnosis of breast lesions by ultrasound: an integrative review. Mastology. 2026;36:e20260013. https://doi.org/10.29289/2594539420260013

Despite these advantages, breast ultrasound is an operator-dependent examination, with prolonged acquisition time and high intra- and inter-observer variability, even when using the Breast Imaging Reporting and Data System (BI-RADS), a standardized system. This limitation contributes to relatively low specificity, increased false positives, more unnecessary biopsies, elevated costs, and a negative impact on patients' experience and quality of life[4–6].

In this context, artificial intelligence (AI), particularly deep learning models, has emerged as a promising tool for mitigating these limitations. AI systems applied to breast ultrasound have demonstrated a potential to support clinical decision-making, functioning as an independent, consistent second reading. The main applications of AI in breast ultrasound include automatic lesion detection and segmentation, diagnostic classification as benign or malignant lesions, and, more recently, prognostic prediction[6–13].

Several studies have shown that AI algorithms can improve the reproducibility of image interpretation, reduce inter-observer variability, and help decrease unnecessary biopsies without compromising diagnostic sensitivity. Thus, integrating breast ultrasound and AI is a significant advance in breast imaging that can positively impact diagnostic accuracy, workflow efficiency, and the standardization of BI-RADS-based reports[6–10].

This integrative review aimed to synthesize and analyze evidence in the literature from comparative studies assessing the performance of AI-assisted classification systems as opposed to conventional (hands-on) breast ultrasound when using BI-RADS, highlighting potential clinical implications and prospects in diagnostic practice.

METHODS

This is an integrative review. A literature search was conducted between January 2016 and January 2024 on PubMed, MEDLINE, LILACS, SciELO, and Connected Papers search platforms, using the following Health Science Descriptors (DeCS) and Medical Subject Headings (MeSH) descriptors: "artificial intelligence," "breast," "ultrasound," "breast neoplasm."

The primary eligibility criterion was the selection of original articles, available online in English, Portuguese, and Spanish, that addressed the use of AI to classify breast lesions, whose images were acquired via conventional (hands-on) ultrasound, regardless of the patient's age, sex, or gender. Another eligibility criterion was that the studies should include images that had undergone histopathological analysis, which is considered the gold standard for statistical analysis of agreement between observers and the AI systems under evaluation. Articles that studied the use of AI in automated systems for breast ultrasound, mammography, tomosynthesis, contrast-enhanced mammography, and magnetic resonance imaging were excluded.

Duplicates, abstracts, letters to journal editors, gray literature, and book chapters were also excluded, as were those that did not mention the topic addressed in this review in their title, abstract, or text. Studies appearing in multiple databases were excluded only after being read in full to avoid exclusion errors.

A single author searched and selected the articles. Subsequently, the same author extracted data from the studies selected using a previously developed and standardized form.

To gain a better understanding of important issues related to the use of AI in ultrasound—and because this is a topic of recent studies—we supplemented the review to identify publications not captured by the electronic search, using secondary references from articles as well as additional searches of the literature on AI in breast ultrasound, employing a broader publication timeframe (extending to February 2026).

Thus, the articles that met all selection criteria and enabled us to address the research question of this review were selected.

In accordance with Resolution No. 466/2012 of the National Health Council (CONEP), this study was exempt from review by a Research Ethics Committee (REC).

RESULTS

Figure 1 details the final selection of 14 articles after evaluation of the inclusion and exclusion criteria.

Of the 14 studies included in the review published between 2016 and 2025, five of them were described as prospective studies[14–18], while nine were presented as retrospective studies[19–27].

The AI technology used in these studies was based on two different systems already approved by the United States Food and Drug Administration and by ANVISA in Brazil. The systems used were: Koios DS v3.1 software (Koios Medical Inc., New York, USA) (ANVISA No. 81464750108) and Samsung S-Detect RS85 Prestige (ANVISA No. 81549259005). The Koios system classifies ultrasound images as BI-RADS 2, BI-RADS 3, BI-RADS 4A/B, and BI-RADS 4C+. The S-Detect system classifies images as likely benign and likely malignant[14–27].

Combining the 14 studies selected for this review, 3,144 patients were evaluated, with a total of 3,524 lesions, all of which underwent histopathological analysis. This resulted in 2,318 benign lesions and 1,206 malignant lesions[14–27].

Patients' ages ranged from 13 to 98 years[14–27]. Bartolotta et al. were the only authors to report the inclusion of men in the study (1.17% of the total number of patients biopsied, corresponding to three nodules in men)[18].

Lesion size ranged from a minimum of 2.0 mm[15] to a maximum of 114 mm[21], measured as the largest diameter of each lesion.

Observers' experience ranged from a minimum of one week[15] to 32 years dedicated to breast radiology[22].

The studies selected for this research had different exclusion criteria, making it difficult to standardize the findings[15,19,23]. For

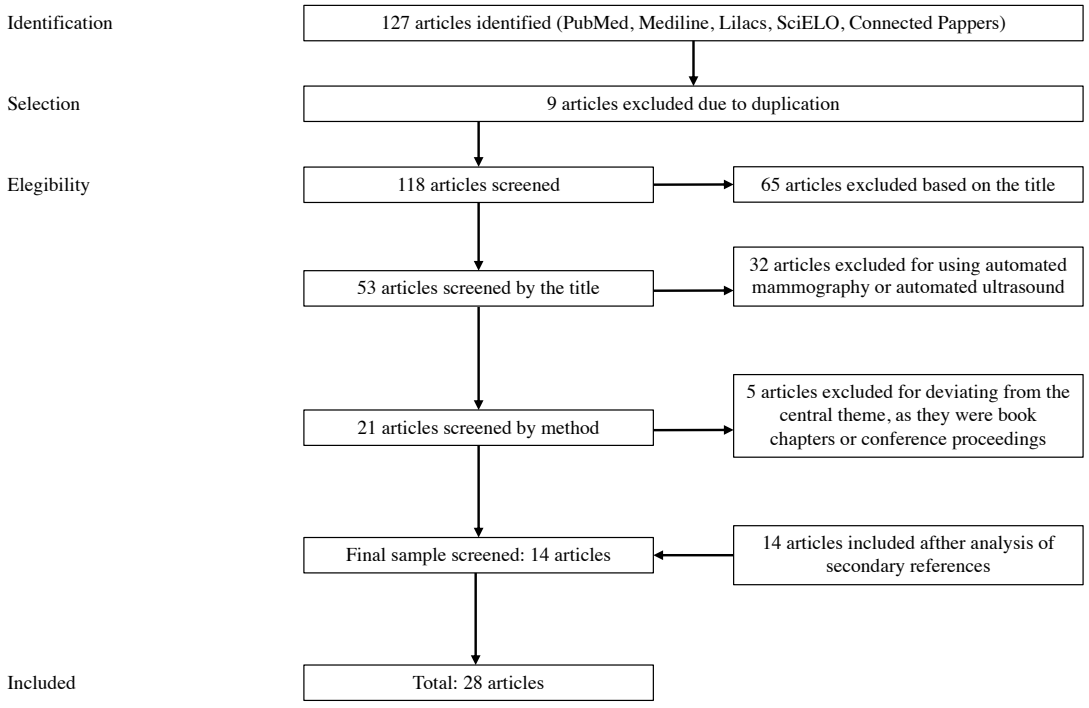


Figure 1. Integrative review flowchart.

example, non-nodular lesions were not included because they were not part of the BI-RADS lexicon at the time. Also, pregnant and/or lactating patients[14,16], patients receiving neoadjuvant chemotherapy[14], and a study of cystic lesions were excluded due to follow-up difficulty[15].

The following table summarizes the 14 studies selected for this integrative review (Table 1).

DISCUSSION

The analyzed studies agree that the use of AI-assisted diagnostic systems represents a useful additional diagnostic tool in breast ultrasound. The benefits, however, vary depending on the practitioner's level of experience and the methodology used to integrate the tool into the workflow[14–27].

A comparative analysis of the studies shows that AI systems applied to breast ultrasound show high and consistent diagnostic performance, particularly in terms of sensitivity and area under the curve (AUC). In general, systems that support diagnostic decision-making (S-Detect and Koios) showed sensitivity ranging from 80% to 95% and AUC ranging from 0.85 to 0.93[14–24,26,27]—values comparable to those reported for experienced radiologists, whose sensitivity ranged from 88% to 96% and AUC from 0.88 to 0.95[14,17,19,20,22,23]. These findings confirm that deep learning-based algorithms can effectively detect malignancies, comparable to that of specialists in breast imaging.

The studies also highlight that AI can effectively distinguish between benign and malignant lesions, reinforcing its potential as a complementary tool in breast ultrasound evaluation. Furthermore, the authors of the studies observed that AI contributed to greater standardization in lesion classification, reducing the subjectivity inherent in ultrasound interpretation and improving diagnostic consistency among different observers[14–20,22–27].

However, specificity had greater variability across studies and, in some scenarios, remained lower than that observed with experienced radiologists when AI was used alone[15,16,19,22,24,26]. Studies conducted by Kim et al.[19], Xing et al.[16], Di Segni et al.[22], Park et al.[23], Guldogan et al.[24], and Wanderley et al.[21] demonstrated specificity ranging from 65% to 82% for automated systems, whereas experienced radiologists frequently achieved values between 75% and 90%. This difference suggests that, although highly sensitive, algorithms may be less effective in reducing false positives compared to experts. This behavior likely reflects the algorithmic design of these tools, which prioritize diagnostic sensitivity to minimize false negatives—a desirable characteristic in cancer screening contexts[16,19,21–24].

A specific analysis of small lesions by Xing et al.[16] demonstrated that, although sensitivity remained high, specificity was slightly reduced in this subgroup. This finding suggests that lesions with less pronounced morphological features may pose a greater challenge for automated algorithms, possibly due to the limited availability of discriminatory features[16].

Furthermore, Heller et al.[25] and Wanderley et al.[21] demonstrated that AI-based decision support systems can significantly contribute to reducing false-positive biopsies, particularly by assisting in the reclassification of benign lesions initially considered

Table 1. List of the 14 studies selected for this integrative review: main results.

Author	Technology	No. of patients	Total lesions	Average diameter	Age (years)	Likely benign	Likely malignant	Experienced observer* (time)	AI	Inexperienced observer ** (time)	AI	AI + Experienced ***	AI + Inexperienced ****
Zhao et al.[14]	S-Detect	227	266 (161/105)		45 (15–82)	2.3	4 A–5	10 years: S: 98.1% E: 65.4% NPV: 98.1% PPV: 64.8% AUC: 0.817 ACUR: 78.2%		2 years: S: 82.8% E: 68.9% NPV: 86.0% PPV: 63.5% AUC: 0.758 ACUR: 74.4%	S: 87.1% E: 72.2% NPV: 81.9% PPV: 79.5% AUC: 0.807 ACUR: 80.4%		
Choi et al.[15]	S-Detect	200	200 (188/12)	9 mm (2–44 mm)	50 (21–77)	2.3	4 A–5	5 years: S: 91.7% E: 76.6% PPV: 20% NPV: 99.3% AUC: 0.84 ACUR: 77%	S: 75% E: 78.2% PPV: 18% NPV: 98% AUC: 0.77 ACUR: 78%	R1 at 1 week: S: 75% E: 71.8% PPV: 14.5% NPV: 97.8% AUC: 0.73 Accuracy: 72%	S: 66.7% E: 76.1% PPV: 15.1% NPV: 97.3% AUC: 0.71 ACUR: 75%	S: 91.7% E: 80.3% PPV: 22% NPV: 99.3% AUC: 0.86 ACUR: 81%	S: 80.3% E: 77.1% PPV: 18.9% NPV: 98.6% AUC: 0.8 ACUR: 77.5%
Xing et al.[16]	S-Detect	295	326 (239/87)	11.9 mm (3–20 mm)	45.4 (18–90)	2–4 A	4 B–5	>10 years: S: 75.9 E: 93.3% PPV: 80.5% NPV: 91.4% AUC: 0.85 ACUR: 88.6%			S: 87.4% E: 88.3% PPV: 73.1% NPV: 98.0% AUC: 0.88 ACUR: 88.4%	S: 90.8% E: 94.6% PPV: 85.9% NPV: 96.6% AUC: 0.93 ACUR: 93.6%	
O'Connell et al.[17]	S-Detect	299	299 (149/150)		52.3	2.3	4 A–5	>10 years: S: 78.1 E: 66.4 AUC: 0.81 ACUR: 71.8	Combined (Exp+inexp) S: 70.3% E: 75.5% AUC: 0.81 ACUR: 73.1%	<5 years S: 62.4 E: 84.6 AUC: 0.8 ACUR: 74.3	S: 80.9 E: 82.7 AUC: 0.84 ACUR: 81.6		
Bartolotta et al.[18]	S-Detect	255	300 (122/178)	13.2 mm (2.6–47 mm)	51 (13–98)	2.3	4.5	>20 years: S: 91.8% E: 81.5% PPV: 77.2% NPV: 93.6% AUC: 0.93				S: 97.5% E: 86.5% PPV: 83.2% NPV: 98.1% AUC: 0.95	
Kim et al.[19]	S-Detect	192	192 (120/72)	14.7 mm (5–52 mm)	46.6 (18–81)	2, 3, 4 A	4 B–5	7 years - breast: S: 76.4% E: 93.3% PPV: 87.3% NPV: 86.8% AUC: 0.772 ACUR: 87%			S: 79.2% E: 65.8% PPV: 58.3% NPV: 84% AUC: 0.725 ACUR: 70.8%		
Xia et al.[20]	S-Detect	40	40 (16/24)		50.9 (± 13.9)	2–4 A	4B -5	>15 years: S: 79.2% E: 81.3% NPV: AUC: 0.802 Accuracy: 60.5%		Junior >3 years: S: 75% E: 68.8% AUC: 0.719 ACUR: 43.8%	S: 95.8% E: 93.8% AUC: 0.948 ACUR: 89.6%	S: 100% E: 93.8% AUC: 0.96	S: 95.8% E: 89.6% AUC: 0.95

It continues...

Table 1. Continuation.

Author	Technology	No. of patients	Total lesions	Average diameter	Age (years)	Likely benign	Likely malignant	Experienced observer* (time)	AI	Inexperienced observer ** (time)	AI	AI + Experienced ***	AI + Inexperienced ****
Wanderley et al.[21]	Koios	509	555 (333/222)	16 mm (3–114 mm)	51 (16–90)	2.3	4A-4B 4C-5	>5 years: S: 99.1% E: 34.0% PPV: 50.1% NPN: 98.3% ACUR 60.2%			S: 98.2% S: 39.0% PPV: 51.8% NPV: 97% ACUR 62.7%		
Di Segni et al.[22]	S-Detect	61	68 (24/44)	10–48 mm	51 (21–84)	2.3	4 A–5	18 and 32 years: S: 93.2% E: 75% PPV: 87.2% NPV: 85.7% AUC: 0.84			S: 91.1% E: 70.8% PPV: 85.4% NPV: 81% AUC: 0.82		
Park et al.[23]	S-Detect	91	100 (59/41)	14 mm (4–39 mm)	46.5 (18–78)	2.3	4.5	8 to 10 years: S: 89.0% E: 53.3% NPV: 87.6% PPV: 57.0% AUC: 0.872 ACUR: 68%		1 year as a trainee S: 76.4% E: 35% NPV: 68.1% PPV: 45.3% AUC: 0.694 ACUR: 51.6%		S: 90.2% E: 66.1% NPV: 90.7% PPV: 64.7% AUC: 0.905 ACUR: 75%	S: 93.5% E: 40.1% NPV: 91.5% PPV: 53.3% AUC: 0.83 ACUR: 62%
Guldogan et al.[24]	Koios	530	715 (581/134)	13.9 mm (3–100 mm)	45 (14–95)	2.3	4.5	5 to 25 years: Radio 1 S: 98.5% E: 80.7% PPV: 54.1% NPV: 99.6% ACUR 84.1%	5 to 25 years: Radio 2 S: 97.8% E: 75.6% PPV: 47.9% NPV: 99.3% ACUR: 79.7%		S: 98.5% E: 65.4% PPV: 39.6% NPV: 99.5% ACUR: 71.6%		
Heller et al.[25]	Koios	200	200 (155/45)			2.3	4.5	6 and 9 years: S: 96.7% E: 61.9% PPV: 42.4% NPV: 98.5% ACUR: 69.8%			S: 100% E: 65.2% PPV: 45.5% NPV: 100% ACUR: 73%		
Çelebi et al.[26]	Koios	70	70 (53/17)	19.9 mm (5–65 mm)	50.4 (32–87)	2.3	4 A–5	15 years: S: 98.1% E: 58.8% PPV: 88.1% NPV: 90.9% AUC: 0.88 ACUR 88.6%		1 year: S: 84.9% E: 17.6% PPV: 76.3% NPV: 27.5% AUC: 0.512 ACUR: 68.6%	S: 92.5% E: 35.3% PPV: 81.7% NPV: 60% AUC: 0.693 ACUR: 78.6%		S: 90.6% E: 17.6% PPV: 77.4% NPV: 37.5% AUC: 0.655 ACUR: 72.9%
Lee et al.[27]	S-Detect	175	193 (120/73)	14.9 mm (3–52 mm)	46.4 (18–81)	2.3	4 A–5	7–19 years: S: 100% E: 30% PPV: 45.4% NPV: 100% AUC: 0.650 ACUR 56.5%			S: 84.9% E: 67.5% PPV: 61.4 NPV: 88.0% AUC: 0.762 Accuracy: 74.1%		

S: Sensitivity; E: Specificity; PPV: Positive predictive value; NPV: Negative predictive value; AUC: Area under the curve; ACUR: Accuracy;
*Interpretation by breast radiologists considered experienced by the authors vs. length of experience in the field.
**Interpretations by breast radiologists considered inexperienced by the authors vs. length of experience in the field.
***Interpretations by breast radiologists considered experienced by the authors, with the aid of AI.
****Interpretations by breast radiologists considered inexperienced by the authors, with the aid of AI.

suspicious. This finding highlights the potential of AI not only as a diagnostic tool but also as a means to optimize clinical decision-making and reduce unnecessary invasive procedures[21,25].

Overall, the studies reviewed have consistent findings regarding the impact of AI on diagnostic performance in breast imaging. The benefit of AI was consistently more pronounced among less experienced radiologists. In this regard, Park et al.[23], Choi et al.[15], and Çelebi et al.[26] showed significant improvement in diagnostic metrics—including sensitivity, specificity, and accuracy—following the incorporation of AI systems, with a particularly significant impact for less experienced readers. Furthermore, these studies also showed reduced inter-observer variability and greater interpretation consistency. Complementarily, Wanderley et al.[21] reinforced these findings by demonstrating that AI contributes to greater diagnostic standardization, consolidating its role as a clinical decision-support tool in radiological practice[15,21,23,26].

Di Segni et al.[22] also observed a significant increase in diagnostic accuracy and improved inter-observer agreement with the use of the S-Detect system, reinforcing the role of these tools in diagnostic standardization. These findings, together with the results described by Wanderley et al.[21], suggest that AI acts as a cognitive support mechanism, reducing the subjectivity inherent in ultrasound, which is an operator-dependent method[21,22].

The negative predictive value was consistently high across the studies, frequently exceeding 90%[17,18,20,21,24]. This finding is particularly relevant to managing lesions classified as BI-RADS 3 and 4A, in which the reliable exclusion of malignancy can prevent unnecessary interventions. Guldogan et al.[24] and Wanderley et al.[21] demonstrated that AI allowed for the reliable reclassification of some of these lesions, suggesting a potential reduction in benign biopsies[21,24].

Another relevant finding concerns the benefit of combining AI with clinical judgment. Studies conducted by Di Segni et al.[22], Park et al.[23], Lee et al.[27], Bartolotta et al.[18], and Wanderley et al.[21] showed that combining radiologists' expertise with AI significantly improved AUC, diagnostic accuracy, and interpretive consistency[18,21–23,27].

Overall, the results show that both systems—S-Detect and Koios—have similar performance, with minor variations attributable to methodological design, study population, and the diagnostic criteria used. These findings reinforce the robustness and reproducibility of these technologies[14,17,19–21].

From a clinical perspective, the findings suggest that AI has promising potential for supporting less experienced radiologists, reducing diagnostic variability and assisting in the reclassification of low-risk lesions[14–27].

Despite promising results, some methodological limitations must be considered. Among the main limitations identified is the retrospective design of most of the studies included[19–27]. Similarly, most studies were conducted at single centers, which may limit the applicability of the results to different clinical settings and populations with distinct epidemiological characteristics[14,19–21,23,26]. This type of study design may introduce selection bias—excluding BI-RADS 2 lesions, for example—since it often includes cases previously selected based on the availability of histopathological confirmation or clinical follow-up. This can overestimate diagnostic performance. Furthermore, many studies included populations with a relatively high prevalence of malignancy, which may overestimate metrics such as positive predictive value and AUC[14–27]. Studies conducted at tertiary centers, such as those by Bartolotta et al.[18], Xia et al.[20], and Wanderley et al.[21], often included a higher proportion of suspicious lesions or those indicated for biopsy, which may positively influence metrics such as sensitivity, positive predictive value, and overall accuracy. Differences in diagnostic reference standards and inclusion criteria also make direct comparisons between studies difficult. In contrast, studies with a higher proportion of benign lesions, such as those by Heller et al.[25] and Lee et al.[27], emphasized the impact of AI in reducing unnecessary biopsies, highlighting its clinical potential in daily practice.

Another relevant limitation concerns the relatively small sample size in some studies[19,22,23,25,26]. Smaller samples may limit statistical robustness and reduce the generalizability of the findings to different clinical populations. Wanderley et al. also highlighted limitations related to sample size and the need for external validation in different populations, emphasizing the importance of prospective multicenter studies to confirm these results[19,22,23,25,26].

In addition, several authors have pointed out key technical limitations, such as the software's inability to assess calcifications within lesions, non-nodular lesions, and vascular flow. The classification of small lesions (diameter ≤1 cm)[20] also poses challenges for the tool[16,18–20,23,28]. Xing et al.[16] further highlight these limitations by noting that the S-Detect AI system appears to focus exclusively on the intrinsic characteristics of the target lesion, neglecting the impact on surrounding tissues, such as discontinuities in the superficial and deep fascia and the presence of architectural distortions—crucial factors in radiological diagnosis. Thus, the need for future technical developments to cover a broader spectrum of ultrasound findings and characterization is highlighted.

In the study by Kim et al.[19], the analysis was based on the fourth edition of the BI-RADS lexicon, which may result in minor discrepancies compared to the newer, more current edition. Furthermore, manual selection of the representative image and delineation of the region of interest by a single experienced radiologist introduced potential operator bias in software interaction[19].

Additionally, technical differences between the systems assessed are another important limitation. The studies included two different AI platforms, S-Detect and Koios DS, with variations in the algorithms used, training methods, and user

interfaces. These technological differences can directly influence diagnostic performance and hinder direct comparisons between studies[17–27].

Critically, the main methodological limitation was that only one of the three authors independently selected the articles. Also, regarding the studies included, we can mention: retrospective study designs, populations from single centers, high prevalence of malignancy, use of static images, and no assessment of long-term clinical outcomes. Thus, we believe future systematic reviews with structured assessments of the methodological quality and risk of bias of the included studies may complement the evidence discussed in this manuscript.

CONCLUSION

Taken together, the available studies demonstrate that AI systems applied to ultrasound exhibit high diagnostic performance, with sensitivity comparable to that of experienced radiologists and moderate to high specificity, particularly when used as a decision-support tool[14–27].

They have also shown benefits, particularly in improving the diagnostic performance of less experienced radiologists, reducing inter-observer variability, and potentially decreasing unnecessary biopsies[14–27].

However, methodological heterogeneity among the studies—including differences in study design, sample size, prevalence of malignancy, and systems used—limits direct comparisons and highlights the need for prospective, multicenter studies with greater methodological standardization[14–27].

Overall, current evidence supports the role of these two AI software programs (S-Detect and the Koios DS system) as promising complementary tools in the ultrasound evaluation of breast lesions, with the potential to improve diagnostic accuracy, standardize interpretations, and optimize clinical management. Future studies are needed to validate these results in larger populations and in more diverse clinical settings, as well as to define the optimal role of these tools in daily clinical practice.

ACKNOWLEDGEMENT

I would like to thank my advisors and colleagues for their support, commitment, and dedication to research into new technologies in medical radiology, a field with a promising future and great potential to benefit patients. We believe that the publication of this study may contribute to improving diagnostic accuracy and optimizing patient care.

AUTHORS' CONTRIBUTIONS

MJGC: Conceptualization, Data curation, Formal Analysis, Project administration, Writing – original draft, Writing – review & editing. BG: Conceptualization, Data curation, Formal Analysis, Visualization. MLL: Conceptualization, Methodology, Writing – review & editing.

Funding: none.

Conflict of interests: nothing to declare.

Artificial Intelligence usage: No generative artificial intelligence tools were used during the preparation of the manuscript.

Data availability statement: The data analyzed in this integrative review were extracted from previously published studies, the results of which are available in the respective articles included in the review. The data compiled and extracted for this analysis may be made available by the corresponding author upon reasonable request.

REFERENCES

1. Brasil. Instituto Nacional de Câncer. Estimativa 2026: incidência de câncer no Brasil. Rio de Janeiro: INCA; 2026; 168 p.
2. 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(3):227-36. https://doi.org/10.1056/nejmoa062790
3. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Böhm-Vélez M, et al. Combined screening with ultrasound and mammography compared to mammography alone in women at elevated risk of breast cancer: results of the first-year screen in ACRIN 6666. JAMA. 2008;299(18):2151-63. https://doi.org/10.1001/jama.299.18.2151
4. Lee SE, Moon JE, Rho YH, Kim EK, Yoon JH. Which supplementary imaging modality should be used for breast ultrasonography? Comparison of the diagnostic performance of elastography and computer-aided diagnosis. Ultrasonography. 2017;36(2):153-9. https://doi.org/10.14366/usg.16033
5. Calas MJG, Alvarenga AV, Gutfilen B, Pereira WCA. Avaliação de parâmetros morfométricos calculados a partir do contorno de lesões de mama em ultrassonografias na distinção das categorias do sistema BI-RADS. Radiol Bras. 2011;44(5):289-96. https://doi.org/10.1590/S0100-39842011000500006
6. Gómez-Flores W, Gregorio-Calas MJ, Pereira WCA. BUS-BRA: A breast ultrasound dataset; for assessing computer-aided

diagnosis systems. Med Phys. 2024;51(4):3110-23. https://doi.org/10.1002/mp.16812

7. Taylor CR, Monga N, Johnson C, Hawley JR, Patel M. Artificial intelligence applications in breast imaging: current status and future directions. Diagnostics (Basel). 2023;13(12):2041. https://doi.org/10.3390/diagnostics13122041
8. Brunetti N, Calabrese M, Martinoli C, Tagliafico AS. Artificial intelligence in breast ultrasound: from diagnosis to prognosis—a rapid review. Diagnostics (Basel). 2023;13(1):58. https://doi.org/10.3390/diagnostics13010058
9. Calas MJG, Lemos ML, Bello MA, Gutfilen B, Bergmann A. Validation of an artificial intelligence program in the characterization of breast nodules by ultrasound. Clin Imaging. 2026;130:110691. https://doi.org/10.1016/j.clinimag.2025.110691
10. Mahant SS, Varma AR. Artificial intelligence in breast ultrasound: the emerging future of modern medicine. Cureus. 2022;14(9):e28945. https://doi.org/10.7759/cureus.28945
11. Liu Q, Qu M, Sun L, Wang H. Accuracy of ultrasonic artificial intelligence in diagnosing benign and malignant breast diseases: a protocol for systematic review and meta-analysis. Medicine (Baltimore). 2021;100(50):e28289. https://doi.org/10.1097/MD.0000000000028289
12. Bahl M, Chang JM, Mullen LA, Berg WA. Artificial intelligence for breast ultrasound: AJR expert panel narrative review. AJR Am J Roentgenol. 2024;223(6):e2330645. https://doi.org/10.2214/ajr.23.30645
13. Brot HF, Mango VL. Artificial intelligence in breast ultrasound: application in clinical practice. Ultrasonography. 2024;43(1):3-14. https://doi.org/10.14366/usg.23116
14. 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;11:921-30. https://doi.org/10.2147/cmar.s190966
15. Choi JH, Kang BJ, Baek JE, Lee HS, Kim SH. Application of computer-aided diagnosis in breast ultrasound interpretation: improvements in diagnostic performance according to reader experience. Ultrasonography. 2018;37(3):217-25. https://doi.org/10.14366/usg.17046
16. Xing B, Chen X, Wang Y, Li S, Liang YK, Wang D. Evaluating breast ultrasound S-detect image analysis for small focal breast lesions. Front Oncol. 2022;12:1030624. https://doi.org/10.3389/fonc.2022.1030624
17. O'Connell AM, Bartolotta TV, Orlando A, Jung SH, Baek J, Parker KJ. Diagnostic performance of an artificial intelligence system in breast ultrasound. J Ultrasound Med. 2022;41(1):97-105. https://doi.org/10.1002/jum.15684
18. 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(7):498-506. https://doi.org/10.1007/s11547-018-0874-7
19. Kim K, Song MK, Kim EK, Yoon JH. Clinical application of S-Detect to breast masses on ultrasonography: a study evaluating the diagnostic performance and agreement with a dedicated breast radiologist. Ultrasonography. 2017;36(1):3-9. https://doi.org/10.14366/usg.16012
20. Xia Q, Cheng Y, Hu J, Huang J, Yu Y, Xie H, et al. Differential diagnosis of breast cancer assisted by S-Detect artificial intelligence system. Math Biosci Eng. 2021;18(4):3680-9. https://doi.org/10.3934/mbe.2021184
21. Wanderley MC, Soares CMA, Morais MMM, Cruz RM, Lima IRM, Chojniak R, et al. Application of artificial intelligence in predicting malignancy risk in breast masses on ultrasound. Radiol Bras. 2023;56(5):329-36. https://doi.org/10.1590/0100-3984.2023.0034
22. Di Segni M, 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(2):105-18. https://doi.org/10.1007/s40477-018-0297-2
23. Park HJ, Kim SM, Yun BL, 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(3):e14146. https://doi.org/10.1097/MD.0000000000014146
24. Guldogan N, Taskin F, Icten GE, Yilmaz E, Turk EB, Erdemli S, et al. Artificial Intelligence in BI-RADS Categorization of Breast Lesions on Ultrasound: Can We Omit Excessive Follow-ups and Biopsies? Acad Radiol. 2024;31(6):2194-202. https://doi.org/10.1016/j.acra.2023.11.031
25. Heller SL, Wegener M, Babb JS, Gao Y. Can an artificial intelligence decision aid decrease false-positive breast biopsies? Ultrasound Q. 2021;37(1):10-15. https://doi.org/10.1097/ruq.0000000000000550
26. Çelebi F, Tuncer O, Oral M, Duymaz T, Orhan T, Ertaş G. Artificial intelligence in diagnostic breast ultrasound: a comparative analysis of decision support among radiologists with various levels of expertise. Eur J Breast Health. 2025;21(1):33-9. https://doi.org/10.1016/j.ultrasmedbio.2023.02.012
27. Lee SE, Moon JE, Rho YH, Kim EK, Yoon JH. Which supplementary imaging modality should be used for breast ultrasonography? Comparison of the diagnostic performance of elastography and computer-aided diagnosis. Ultrasonography. 2017;36(2):153-9. https://doi.org/10.14366/usg.16033
28. 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. https://doi.org/10.1016/j.ejrad.2021.109624

This is an open access article distributed under the terms of the Creative Commons license.