Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Diagnostic Accuracy of Conventional Breast Ultrasound in Differentiating Benign and Malignant Breast Lesions: A Histopathological Correlation Study at a Tertiary Military Hospital in Rwanda (2023–2025)

Aime Ishimwe Mugisha
University of Rwanda
James Nelson Okema
Gulu University
Egide Ndayambaje
University of Rwanda
Gerald Olwit
geraldolwit@gmail.com

Kabale University
Arnold Atuhaire
Kabale University
Emmanuel Kubana
University of Rwanda
Linda Kapesa
Butaro Hospital
Jackson Muvunyi
Rwanda Military Hospital
Eric Seruyange
Rwanda Military Hospital
Emile Karinganire
Rwanda Military Hospital
Emmanuel Rudakemwa
Rwanda Military Hospital

Research Article

Keywords: breast ultrasound, BI-RADS, histopathology, diagnostic accuracy, breast cancer, sub-Saharan Africa, Rwanda

Posted Date: November 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10614848/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Breast cancer is the most commonly diagnosed cancer among women worldwide, and outcomes in sub-Saharan Africa remain poor, largely because of late-stage presentation and limited diagnostic infrastructure. Conventional ultrasound, reported using the Breast Imaging Reporting and Data System (BI-RADS), is the primary breast imaging tool in many low-resource settings, yet locally generated evidence on its diagnostic performance is scarce in Rwanda.

Methods

We conducted a retrospective cross-sectional diagnostic accuracy study of 500 patients who underwent conventional breast ultrasound with documented BI-RADS classification and subsequent histopathological evaluation at Rwanda Military Teaching Hospital between January 2023 and December 2025. BI-RADS categories were dichotomized (2–3 as negative, 4–5 as positive for malignancy) and compared with histopathological diagnosis, the reference standard. Sensitivity, specificity, predictive values, and overall accuracy were calculated from a 2×2 table; agreement was assessed with Cohen's kappa; the discriminatory performance of ordinal BI-RADS category was assessed with receiver operating characteristic (ROC) curve analysis.

Results

The mean age was 44.6 ± 16.6 years and 97.8% of patients were female. Histopathology confirmed malignancy in 329 lesions (65.8%), most commonly invasive ductal carcinoma – no special type (58.6% of all lesions), and benign disease in 171 lesions (34.2%), most commonly fibroadenoma (14.6%). Ultrasound correctly identified 271 of 329 malignant lesions and 117 of 171 benign lesions, giving a sensitivity of 82.4% (95% CI 77.9–86.1), specificity of 68.4% (95% CI 61.1–74.9), positive predictive value of 83.4% (95% CI 79.0–87.0), negative predictive value of 66.9% (95% CI 59.6–73.4), and overall accuracy of 77.6% (95% CI 73.7–81.0). Agreement between BI-RADS classification and histopathology was moderate ($\kappa = 0.505$, 95% CI 0.425–0.585, $p < 0.001$), and BI-RADS category was significantly associated with histopathological diagnosis ($\chi^2 = 127.594$, df = 1, $p < 0.001$). ROC analysis showed good discrimination (AUC = 0.819, 95% CI 0.781–0.857, $p < 0.001$).

Conclusion

In this, to our knowledge, largest Rwandan series correlating breast ultrasound with histopathology, conventional ultrasound showed good sensitivity but only moderate specificity and agreement with histopathology, with a false-negative rate of nearly 18% among histologically malignant lesions.

Ultrasound alone should not be used to exclude malignancy in clinically or sonographically indeterminate breast masses; a low threshold for biopsy is warranted in this setting.

Introduction

Breast lesions span a heterogeneous spectrum, from benign proliferative changes such as fibroadenomas, cysts, and fibrocystic disease to invasive malignant disease, and benign and malignant lesions frequently share overlapping clinical and sonographic features (1,2). Breast cancer remains the most commonly diagnosed cancer among women worldwide and a leading cause of cancer-related morbidity and mortality (3,4). Its burden has shifted increasingly toward low- and middle-income countries (LMICs), where rising incidence coexists with persistently high mortality driven by late-stage presentation, limited screening access, inadequate diagnostic infrastructure, and delayed treatment initiation (5–7). Notably, screening-uptake barriers are not unique to LMIC health systems: comparable obstacles relating to awareness and health-system navigation have been documented among women of Black African descent resident in high-income countries (22,23), reinforcing that late-stage presentation reflects systemic access constraints rather than a phenomenon confined to resource-poor infrastructure alone. In sub-Saharan Africa, including Rwanda, these challenges are compounded by resource constraints, limited imaging availability outside referral centres, and variable diagnostic expertise, all of which contribute to poorer outcomes than in high-income settings.

Ultrasound is one of the most widely used breast imaging modalities because of its accessibility, affordability, lack of ionizing radiation, and utility in dense breast tissue (8,9), and in many low-resource settings it functions as the primary, rather than adjunctive, breast imaging tool. The Breast Imaging Reporting and Data System (BI-RADS) standardizes interpretation and reporting, assigning an estimated risk of malignancy to each finding and guiding subsequent management (10,11). Histopathological examination remains the reference standard for definitive diagnosis (12). Reported sensitivity of breast ultrasound for malignancy ranges from approximately 80% to 98%, while specificity is more variable, from 60% to 95%, reflecting operator dependency and differences in lesion characteristics and patient populations (13–15). More advanced modalities, such as contrast-enhanced spectral mammography and breast MRI, have shown higher diagnostic accuracy than conventional imaging in meta-analytic comparisons (18), but their cost and infrastructure requirements limit their availability in most low-resource settings, reinforcing the continued centrality of conventional ultrasound as the primary breast imaging tool in settings such as Rwanda.

Much of the existing literature has focused on malignant lesions in relatively small, selected samples, limiting generalizability to routine practice where both benign and malignant lesions are encountered (17,24). Evidence from sub-Saharan Africa is particularly limited, and available data suggest that while sensitivity is generally preserved, specificity varies considerably with equipment, operator training, disease stage at presentation, and health-system factors (24,25). In Rwanda specifically, published diagnostic accuracy data are scarce; most available information derives from descriptive hospital-based reports rather than systematic correlation with histopathology.

At Rwanda Military Teaching Hospital (RMTH), breast ultrasound is performed routinely in patients presenting with breast abnormalities, and many subsequently undergo histopathological evaluation. This study used this existing clinical dataset to determine, for the first time in this setting, the diagnostic accuracy of conventional breast ultrasound against histopathology and the agreement between the conventional ultrasound and histopathology in diagnosing breast lesion in an unselected population that included both benign and malignant lesions. We hypothesized a priori that, consistent with prior reports from other resource-limited settings, conventional ultrasound would demonstrate high sensitivity for malignancy but comparatively lower specificity in this unselected population.

Methods

Study design and setting

This was a retrospective, cross-sectional diagnostic accuracy study conducted at Rwanda Military Teaching Hospital (RMTH), a tertiary referral and teaching hospital in Kigali, Rwanda, that provides comprehensive breast imaging, pathology, surgical, and oncology services to patients from across the country. This study is reported in accordance with the Standards for Reporting of Diagnostic Accuracy Studies (STARD 2015) guideline; the completed STARD checklist is provided as a supplementary file, and the patient flow diagram (Fig. 1) follows the STARD-recommended format.

Study population, eligibility, and sampling

The study population comprised patients evaluated for breast abnormalities at Rwanda Military Teaching Hospital (RMTH), a tertiary referral hospital in Kigali, Rwanda, between January 2023 and December 2025. Patient records were retrospectively identified from the hospital electronic medical record (EMR) system, with relevant radiology and histopathology records reviewed for diagnostic correlation. Data were extracted using a predefined Excel data-extraction sheet containing coded variables of interest.

The patient-selection process is summarized in Fig. 1. The initial review identified 2,762 breast ultrasound examinations documenting breast abnormalities during the study period. Of these, 889 records were excluded because the ultrasound reports did not contain a documented BI-RADS classification, leaving 1,873 examinations with a documented BI-RADS assessment. Among these, 634 patients had been referred for histopathological evaluation for further diagnostic assessment. The remaining 1,239 patients had non-complicated lesions that did not require histopathological confirmation and were therefore not eligible for inclusion in the present diagnostic accuracy analysis. Of the 634 patients referred for histopathological evaluation, 83 were excluded because the histopathological reports could not be matched to the corresponding BI-RADS ultrasound records, while a further 51 were excluded because the relevant histopathology reports were unavailable in the medical records. The final study population comprised 500 patients with documented BI-RADS classification and corresponding histopathological diagnoses. Within this final cohort, all 500 index-test and reference-

standard results were categorical and interpretable; no ultrasound report lacked a definitive BI-RADS assignment and no histopathology report was equivocal or deferred, so no additional handling rule for indeterminate results was required beyond the exclusions already described above.

Data collection, index test, and reference standard

Data were retrospectively extracted from the hospital electronic medical record system, radiology and histopathology reports, and relevant hospital databases using a predefined Excel data-extraction sheet with coded variables of interest. Extracted variables included demographic characteristics, ultrasound findings (BI-RADS category, laterality, and lesion characteristics where available), and histopathological diagnosis, including benign or malignant classification and specimen type. Histopathological specimens included core needle, excisional, incisional, and mastectomy specimens, among others.

Conventional breast ultrasonography, interpreted using the BI-RADS classification system, was the index test, while histopathological diagnosis served as the reference standard. Histopathology was selected as the reference standard because it provides direct tissue diagnosis and is the accepted gold standard against which breast imaging classification systems such as BI-RADS are conventionally validated. Ultrasound examinations were performed and reported by two radiology officers with formal medical-imaging training and additional specialized training in ultrasonography, and two consultant radiologists; the specific ultrasound equipment make and model, transducer frequency, and step-by-step scanning protocol were not uniformly documented in the retrospective imaging records and could not be extracted for this analysis, but examinations were performed according to the standard departmental breast-ultrasound protocol in use at RMTH throughout the study period. BI-RADS categories 2–3 was classified as negative and categories 4–5 as positive for malignancy, a dichotomization that reflects standard, widely used clinical convention and was pre-specified before data analysis began rather than derived post hoc from the study data. Histopathological diagnoses were extracted from the final pathology reports and classified as benign or malignant; the number and training level of the reporting pathologists and the specific tissue-processing and reporting standards applied (e.g., use of synoptic reporting or immunohistochemistry) were not consistently recorded in the retrospective pathology reports and could not be systematically extracted for this analysis. The interval between ultrasound examination and histopathological sampling was likewise not uniformly recorded and could not be systematically reported; no clinical intervention intended to alter the breast lesion occurred between the index test and the reference standard in any patient.

The patient was the unit of analysis. For patients with bilateral lesions, the lesion corresponding to the histopathological specimen used for diagnostic confirmation was analyzed, ensuring one observation per patient. Because radiology and pathology records were accessible through the shared hospital electronic medical record system, formal blinding of radiologists to histopathological findings and pathologists to ultrasound/BI-RADS findings could not be confirmed retrospectively.

Statistical analysis

No a priori sample size calculation was performed; the study included all consecutive patients meeting eligibility criteria during the study period (a consecutive/convenience sample), and the final sample of 500 patients was determined by the number of eligible records identified through the process shown in Fig. 1 rather than by a pre-specified target. Data were entered, cleaned, and analysed using IBM SPSS Statistics version 20 (IBM Corp., Armonk, NY, USA). Continuous variables were summarized as means and standard deviations; categorical variables as frequencies and percentages. BI-RADS categories were dichotomized into negative (2–3) and positive/suspicious (4–5) groups, and histopathological diagnoses were classified as benign or malignant, to construct a 2×2 contingency table. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall diagnostic accuracy were calculated using histopathology as the reference standard, with 95% confidence intervals for sensitivity, specificity, and predictive values computed using the Wilson score method; the 95% confidence interval for the AUC was computed using the asymptotic nonparametric method (Hanley and McNeil) as implemented in SPSS's ROC Curve procedure. The association between BI-RADS classification and histopathological diagnosis was assessed with Pearson's chi-square test. Agreement was quantified using Cohen's kappa (κ), interpreted as moderate (0.41–0.60), substantial (0.61–0.80), or almost perfect (> 0.80). Receiver operating characteristic (ROC) curve analysis was performed to assess discriminatory performance, with the area under the curve (AUC) and 95% confidence interval (CI) reported. This analysis used the four-level ordinal BI-RADS category (2, 3, 4, 5) as the test variable against the binary histopathological outcome, distinct from the dichotomized BI-RADS 4–5 versus 2–3 cut-off used for sensitivity, specificity, and predictive value calculations; because the test variable has only four distinct levels, the resulting ROC curve is a step function with three non-trivial operating points rather than a smooth curve. A two-sided p-value < 0.05 was considered statistically significant. The primary analysis was restricted to patients with complete BI-RADS and histopathological classification data (complete-case analysis); patients with incomplete lesion descriptors beyond these two core variables were retained in the primary 2×2 analysis but excluded from any secondary analyses using those descriptors. No pre-specified or exploratory subgroup, stratified, or sensitivity analyses (for example, by sex, specimen type, or BI-RADS 4 sub-category) were conducted for the primary diagnostic-accuracy estimates reported here; such analyses are identified as recommended future work in the Limitations section.

Ethical considerations

Ethical approval

for this retrospective study was obtained from the Rwanda Military Teaching Hospital Institutional Review Board (RMTH/IRB; reference number 321/RMTH/COMDT/2026; approved 14 July 2026). The study involved secondary analysis of routinely collected clinical, radiological, and histopathological records from January 2023 to December 2025; no prospective patient recruitment or new data collection was undertaken. Because the study used existing records and involved no direct patient contact, the requirement for individual informed consent was waived by the Institutional Review Board. All data were

anonymized before analysis, and the study was conducted in accordance with the principles of the Declaration of Helsinki. Ultrasound, the index test, is non-invasive and carries no procedural risk; histopathological sampling (the reference standard) is an invasive procedure with recognised standard clinical risks, but because this was a retrospective review of existing records rather than a prospectively monitored study, adverse-event data related to either the index test or the reference standard were not captured and are not reported.

The results will be presented to the RMTH/IRB prior to publication, in accordance with the terms of the ethical clearance.

Results

Baseline characteristics

Five hundred patients with breast mass lesions were included. Mean age was 44.6 ± 16.6 years; 25.4% were aged 40–49 years, 20.4% were younger than 30 years, 18.8% were aged 50–59 years, 18.6% were 60 years or older, and 16.8% were aged 30–39 years. Females comprised 97.8% of the cohort and males 2.2%. Lesions were right-sided in 52.6%, left-sided in 46.0%, and bilateral in 1.4%. The tissue diagnosis was obtained by core needle biopsy in 33.4% of cases, excisional biopsy in 29.2%, lumpectomy in 25.8%, mastectomy specimen in 10.2%, and incisional biopsy in 1.4% (Table 1).

Table 1
Baseline demographic and clinical characteristics (N = 500)

Variable	***n* (%)**
Age, years, mean ± SD	44.6 ± 16.6
<30	102 (20.4)
30–39	84 (16.8)
40–49	127 (25.4)
50–59	94 (18.8)
≥60	93 (18.6)
Sex: Male	11 (2.2)
Sex: Female	489 (97.8)
Laterality: Right	263 (52.6)
Laterality: Left	230 (46.0)
Laterality: Bilateral	7 (1.4)
Specimen: Core needle biopsy	167 (33.4)
Specimen: Excisional biopsy	146 (29.2)
Specimen: Incisional biopsy	7 (1.4)
Specimen: Lumpectomy	129 (25.8)
Specimen: Mastectomy specimen	51 (10.2)

Spectrum of histopathological and sonographic diagnoses

Invasive ductal carcinoma – no special type (IDC-NST) was the single most common histopathological diagnosis, accounting for 293 of 500 lesions (58.6%), followed by fibroadenoma (73, 14.6%) and fibrocystic disease (43, 8.6%). Invasive lobular carcinoma accounted for 17 cases (3.4%), tubular adenoma for 15 (3.0%), and benign phyllodes tumour for 10 (2.0%); the remaining diagnoses each accounted for ≤ 1.6% of cases (Table 2). On ultrasound, solid mass was the most frequent descriptor (218, 43.6%), followed by fibroadenoma (90, 18.0%) and non-specific breast mass (84, 16.8%); suspicious mass was reported in 34 cases (6.8%) and a normal examination in 21 (4.2%) (Table 3). Histological subtype distribution among malignant lesions, presented in Table 2, is reported as a proxy indicator of disease spectrum; formal measures of disease severity such as tumour size, TNM stage, or histological grade were not consistently captured in the retrospective pathology records and are therefore not reported.

Table 2
Spectrum of histopathological diagnoses among breast mass lesions (N = 500)

Histopathological diagnosis	***n***	**%**
Invasive ductal carcinoma – NST (IDC-NST)	293	58.6
Fibroadenoma	73	14.6
Fibrocystic disease	43	8.6
Invasive lobular carcinoma	17	3.4
Tubular adenoma	15	3.0
Benign phyllodes tumour	10	2.0
Other	8	1.6
In situ carcinoma	5	1.0
Lipoma	5	1.0
Lactating adenoma	4	0.8
Malignant phyllodes tumour	4	0.8
Papillary carcinoma	4	0.8
Atypical ductal hyperplasia (ADH)	3	0.6
Gynecomastia	2	0.4
Malignant mesenchymal neoplasm	2	0.4
Metaplastic carcinoma	2	0.4
Mucinous carcinoma	2	0.4
Cyst / dermoid cyst / ductal hyperplasia / hypertrophic scar / intraductal papilloma / mammary duct ectasia / morphea / normal histology (each)	1	0.2
Total	500	100.0

Table 3
Spectrum of ultrasonographic findings among breast mass lesions (N = 500)

Ultrasound diagnosis	***n***	**%**
Solid mass	218	43.6
Fibroadenoma	90	18.0
Breast mass (non-specific)	84	16.8
Suspicious mass	34	6.8
Normal breast	21	4.2
Complex mass	10	2.0
Cyst	9	1.8
Breast nodule	7	1.4
Complex cystic mass	7	1.4
Fibrocystic changes	5	1.0
Lipoma	3	0.6
Ductal cell carcinoma	2	0.4
Phyllodes tumour	2	0.4
Calcified mass / breast abscess / dermoid cyst / DCIS / gynecomastia / lymph node / mammary duct ectasia / sebaceous cyst (each)	1	0.2
Total	500	100.0

Association and agreement between BI-RADS classification and histopathology

Among 329 histopathologically malignant lesions, 271 (82.4%) were classified as malignant/suspicious (BI-RADS 4–5) and 58 (17.6%) as benign (BI-RADS 2–3) on ultrasound. Among 171 histopathologically benign lesions, 117 (68.4%) were classified as benign and 54 (31.6%) as malignant/suspicious on ultrasound. Overall, 388 of 500 lesions (77.6%) showed concordant classification. Pearson's chi-square analysis confirmed a statistically significant association between BI-RADS classification and histopathological diagnosis ($\chi^2 = 127.594$, df = 1, $p < 0.001$) **(**Table 4**)**.

Table 4
Association between ultrasonographic BI-RADS classification and histopathological diagnosis

Histopathological diagnosis	BI-RADS benign (2–3)	BI-RADS malignant/suspicious (4–5)	Total
Malignant	58 (17.6%)	271 (82.4%)	329
Benign	117 (68.4%)	54 (31.6%)	171
Total	175	325	500

$\chi^2 = 127.594$, $df = 1$, $p < 0.001$.

Agreement between BI-RADS classification and histopathological diagnosis, assessed with Cohen's kappa, was statistically significant and moderate in magnitude ($\kappa = 0.505$, SE = 0.041, $p < 0.001$), consistent with the 77.6% observed agreement (Table 5).

Table 5
Agreement between ultrasonographic BI-RADS classification and histopathological diagnosis

Measure	Value
Number of lesions	500
Observed agreement (%)	77.6
Cohen's kappa (κ)	0.505
Standard error	0.041
95% confidence interval	0.425–0.585
p-value	< 0.001

Diagnostic performance of ultrasonography

Using histopathology as the reference standard, ultrasonography correctly identified 271 of 329 malignant lesions and 117 of 171 benign lesions. Sensitivity was 82.4% (95% CI 77.9–86.1), specificity 68.4% (95% CI 61.1–74.9), positive predictive value 83.4% (95% CI 79.0–87.0), negative predictive value 66.9% (95% CI 59.6–73.4), and overall diagnostic accuracy 77.6% (95% CI 73.7–81.0) (Table 6). Cohen's kappa was 0.505 (95% CI 0.425–0.585). ROC curve analysis of the ordinal BI-RADS category showed good discriminatory performance for distinguishing malignant from benign lesions, with an AUC of 0.819 (95% CI 0.781–0.857, $p < 0.001$) (Fig. 2), indicating that the probability of malignancy rose consistently with increasing BI-RADS category.

Table 6
Diagnostic performance of ultrasonography for detecting malignant breast lesions

Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	Accuracy, % (95% CI)	AUC (95% CI)
82.4 (77.9–86.1)	68.4 (61.1–74.9)	83.4 (79.0–87.0)	66.9 (59.6–73.4)	77.6 (73.7–81.0)	0.819 (0.781–0.857)

Discussion

In this retrospective correlation study of 500 patients evaluated with both conventional breast ultrasound and histopathology at a Rwandan tertiary hospital, ultrasound demonstrated good sensitivity (82.4%) but only moderate specificity (68.4%) for malignancy, with moderate overall agreement with histopathology (κ = 0.505) and good discriminatory performance across BI-RADS categories (AUC = 0.819). These findings both confirm and extend the existing literature: the sensitivity observed here sits within the range reported by prior diagnostic accuracy studies (approximately 80–98%) (13–15), while the specificity is at the lower end of the previously reported 60–95% range, likely reflecting our inclusion of an unselected, real-world population with both benign and malignant lesions rather than a malignancy-enriched sample (24,25).

Two findings merit particular clinical attention. First, malignancy was unusually prevalent in this cohort, 65.8% of lesions were histologically malignant, with IDC-NST alone accounting for nearly six in ten specimens. This most plausibly reflects the referral and biopsy pathway at a tertiary centre performing tissue sampling on clinically or sonographically concerning lesions, rather than the true population prevalence of malignancy among all women presenting with breast symptoms, and should be considered when generalizing these results to lower-acuity or screening settings. Second, and most clinically important, 58 of 329 histologically malignant lesions (17.6%) were classified as BI-RADS 2–3 (benign) on ultrasound. This false-negative proportion is consistent with reports from other sub-Saharan African settings attributing lower specificity and higher miss rates to operator dependency, equipment variability, and a higher proportion of advanced or atypically presenting disease at diagnosis (22,23). It reinforces that a benign sonographic impression cannot reliably exclude malignancy in this population, and that biopsy thresholds should remain low for masses that are clinically suspicious even when BI-RADS category is reassuring.

The moderate kappa statistic, despite a statistically significant chi-square association and good AUC, illustrates an important methodological point relevant to diagnostic accuracy reporting: a strong ordinal relationship between BI-RADS category and probability of malignancy (as captured by the ROC/AUC) can coexist with only moderate categorical agreement once BI-RADS is dichotomized at a single operating threshold (4 vs 3). This suggests that refinement of the BI-RADS 3/4A boundary, or closer attention to the intermediate categories, may improve the practical accuracy of the dichotomized clinical decision

rule in this setting, consistent with observations elsewhere that diagnostic performance varies across BI-RADS sub-categories rather than being uniform (20,21).

These results should be interpreted alongside those of other sub-Saharan African series. Prior regional studies correlating sonographic and histopathological findings in breast masses have similarly reported preserved sensitivity with more variable specificity (17,24,25), and reviews of breast imaging and pathology capacity in the region have highlighted resource and training constraints as contributors to this variability (19). To our knowledge, this is among the first studies from Rwanda to report the full diagnostic accuracy profile, sensitivity, specificity, predictive values, kappa agreement, and ROC/AUC of conventional breast ultrasound against histopathology in a cohort that included both benign and malignant lesions, addressing a gap noted in the regional literature (17,24,25).

Strengths and limitations

This study benefits from a relatively large, consecutively sampled cohort for a single-centre diagnostic accuracy study in this setting, the inclusion of both benign and malignant lesions rather than malignancy-enriched samples, and the use of a comprehensive set of accuracy and agreement metrics recommended for diagnostic accuracy reporting. Several limitations should be considered. The retrospective design relies on the completeness and quality of existing radiology and pathology records and may be subject to missing or inconsistent data. The reference standard itself was heterogeneous, comprising core needle, incisional, excisional, and mastectomy specimens; core needle biopsy in particular carries a recognised sampling-error false-negative rate for malignancy, so some lesions classified as benign by the reference standard may represent undersampled malignancies, which would bias specificity and NPV upward rather than downward and should be considered when interpreting these estimates. Because only patients who underwent both ultrasound and histopathology were eligible, the cohort is subject to verification (work-up) bias: patients whose ultrasound was reassuring and who were not referred for biopsy are not represented, which is expected to inflate the apparent sensitivity and deflate specificity relative to a true screening population; the very high malignancy prevalence in this cohort (65.8%) relative to unselected symptomatic populations is itself evidence that this bias is operating, and the reported accuracy figures should therefore be read as an upper bound for sensitivity and a corresponding underestimate of real-world specificity, rather than as population-level screening performance. Relatedly, the demographic and clinical characteristics of the 134 patients excluded from the histopathology-referred group because their records could not be matched or retrieved (Fig. 1) were not available in the dataset used for this analysis, so a formal comparison with the 500 included patients could not be performed; the direction and magnitude of any additional selection bias this introduces cannot be determined from the data at hand, and this should be read as compounding, rather than replacing, the verification bias described above. Atypical ductal hyperplasia (n = 3) was classified as benign for the purposes of the reference standard; as a recognised risk lesion with a non-trivial upgrade rate to malignancy on excision, this is a simplification that should be acknowledged. Patient-level linkage between histological subtype and BI-RADS category was not retained in the retrospective extraction used for this analysis, precluding an exact recalculation excluding

these cases; a worst-case bounding analysis was therefore performed instead, considering both extremes of how these 3 cases could have been classified on ultrasound. Because ADH falls entirely within the benign reference-standard group, sensitivity is mathematically unaffected (82.4%); across the two extremes, specificity ranges from 67.9% to 69.6% (base case 68.4%), negative predictive value from 66.3% to 66.9% (base case 66.9%), and overall accuracy from 77.5% to 78.1% (base case 77.6%). This narrow range indicates that the classification of these 3 cases cannot materially alter the study's principal findings under any allocation. Eleven male patients (2.2%) were included; because male breast disease differs biologically and radiologically from female breast disease, a sex-restricted sensitivity analysis would be the preferred approach, but sex was not linked to individual index-test and reference-standard outcomes in the retrospective extraction, precluding an exact recalculation. A worst-case bounding analysis was performed instead, considering the extremes in which all 11 male patients fell in the direction least, versus most, favourable to each metric: sensitivity ranges from 81.8% to 85.2% (base case 82.4%), specificity from 66.2% to 73.1% (base case 68.4%), and overall accuracy from 77.1% to 79.3% (base case 77.6%). This range is wider than for the ADH subgroup, reflecting the larger subgroup size, and specificity in particular could plausibly shift by several percentage points depending on the true classification of these patients; nonetheless, even the lower bound for each metric remains consistent with the qualitative conclusions reported here (good sensitivity, moderate specificity), so the inclusion of male patients is unlikely to reverse the study's principal findings, though this bound should temper confidence in generalising the point estimates to a female-only population. The eight lesions (1.6%) grouped as "Other" in Table 2 reflect a residual free-text category in the retrospective pathology records that was not coded to individual diagnosis at the point of data extraction; the original records could not be re-accessed to retrieve this detail for the present analysis, so this category should be interpreted as histologically heterogeneous rather than as a single entity. As a single-centre study at a tertiary referral hospital, findings may not generalize to primary or secondary-level facilities with different equipment or operator experience. Radiologist-level variation in interpretation could not be assessed, as operator identity was not captured. Finally, lesion-level sonographic descriptors beyond BI-RADS category (e.g., margin, shape, orientation) were incompletely available and could not be incorporated into a multivariable model of diagnostic performance.

Conclusion

Conventional breast ultrasound at this Rwandan tertiary hospital showed good sensitivity and good overall discriminatory performance for malignancy but only moderate specificity and moderate agreement with histopathology, with a clinically important false-negative rate of nearly 18% among histologically malignant lesions. These findings support continued reliance on ultrasound as a first-line, accessible breast imaging tool in this setting, while underscoring that a benign BI-RADS assignment should not, by itself, be used to defer biopsy in clinically suspicious masses. Prospective, multicentre studies incorporating operator-level data and full BI-RADS sub-categorization are needed to refine locally applicable diagnostic thresholds and to determine whether targeted radiologist training could narrow the specificity gap identified here.

Declarations

Ethics approval and consent to participate: This study was approved by the Rwanda Military Teaching Hospital Institutional Review Board (reference 321/RMTH/COMDT/2026, approved 14 July 2026). As a retrospective review of anonymized existing records, the IRB granted a waiver of individual informed consent.

Consent for publication: Not applicable.

Trial or study registration: This retrospective diagnostic accuracy study was not registered in a clinical trial or study registry.

Study protocol: No separate study protocol was prepared or registered in advance, consistent with the retrospective, record-based design of this study; the methods reported in full in this manuscript constitute the complete description of study conduct.

Availability of data and materials: The datasets analysed during the current study are available from the corresponding author on reasonable request, subject to institutional data-sharing policies.

Competing interests: The authors declare no competing interests.

Funding: This study received no external funding from any funding agency in the public, commercial, or not-for-profit sectors

Authors' contributions: AIM, EN, KE, LK, JM, ES, AA, EK, and ER contributed to data acquisition. AIM, JNO, and GO performed the statistical analysis and interpretation of data. AIM and JNO drafted the manuscript. All authors (AIM, JNO, EN, GO, KE, LK, JM, ES, AA, EK, ER) critically revised the manuscript for important intellectual content, approved the final version, and agree to be accountable for all aspects of the work.

Acknowledgements: The authors thank the Departments of Radiology and Pathology, Rwanda Military Teaching Hospital, for facilitating access to records.

References

1. GBD 2023 Breast Cancer Collaborators. Global, regional, and national burden of breast cancer among females, 1990–2023, with forecasts to 2050: a systematic analysis for the Global Burden of Disease Study 2023. Lancet Oncol. 2026;27(3):302–26. doi:10.1016/S1470-2045(25)00730-2
2. Arzanova E, Mayrovitz HN. The epidemiology of breast cancer. In: Mayrovitz HN, editor. Breast Cancer. Brisbane: Exon Publications; 2022. p. 1–20. doi:10.36255/exon-publications-breast-cancer-epidemiology
3. Freihat O, Sipos D, Kovacs A. Global burden and projections of breast cancer incidence and mortality to 2050: a comprehensive analysis of GLOBOCAN data. Front Public Health. 2025; 13:1622954.

doi:10.3389/fpubh.2025.1622954

4. Kim J, Harper A, McCormack V, Sung H, Houssami N, Morgan E, et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries. Nat Med. 2025;31(4):1154–62. doi:10.1038/s41591-025-03502-3
5. Tolentino-Rodriguez L, Chkeir M, Pofagi V, Ahindu I, Toniolo J, Erazo A, et al. Breast cancer characteristics in low- and middle-income countries: an umbrella review. Cancer Epidemiol. 2025; 96:102797. doi: 10.1016/j.canep.2025.102797
6. Pace LE, Shulman LN. Breast cancer in sub-Saharan Africa: challenges and opportunities to reduce mortality. Oncologist. 2016;21(6):739–44. doi:10.1634/theoncologist.2015-0429
7. Jedy-Agba E, McCormack V, Adebamowo C, dos-Santos-Silva I. Stage at diagnosis of breast cancer in sub-Saharan Africa: a systematic review and meta-analysis. Lancet Glob Health. 2016;4(12): e923–35. doi:10.1016/S2214-109X (16)30259-5
8. Dan Q, Zheng T, Liu L, Sun D, Chen Y. Ultrasound for breast cancer screening in resource-limited settings: current practice and future directions. Cancers (Basel). 2023;15(7):2112. doi:10.3390/cancers15072112
9. Evans A, Trimboli RM, Athanasiou A, Balleyguier C, Baltzer PA, et al., for the European Society of Breast Imaging (EUSOBI). Breast ultrasound: recommendations for information to women and referring physicians by the European Society of Breast Imaging. Insights Imaging. 2018;9(4):449–61. doi:10.1007/s13244-018-0636-z
10. Merjane V, Perin DMP, Bacha PMGE, Miranda BMM, Bitencourt AGV, Iared W. Breast Imaging Reporting and Data System (BI-RADS®): a success history and particularities of its use in Brazil. Rev Bras Ginecol Obstet. 2024;46: e-rbgo6. doi:10.61622/rbgo/2024AR06
11. Minichetti P, Cereser L, Versienti E, Francioso F, Sparascio F, Girometti R, et al. BI-RADS v2025: key updates and implications for breast imaging practice. AJR Am J Roentgenol. 2026; AJR.26.35051. doi:10.2214/AJR.26.35051
12. Alotaibi BS, Alghamdi R, Aljaman S, Hariri RA, Althunayyan LS, AlSenan BF, et al. The accuracy of breast cancer diagnostic tools. Cureus. 2024;16(1): e51776. doi:10.7759/cureus.51776
13. Rehman H, Ahmad I, Rashid S, Mukhtar M, Khan AA, Khaliq H. Comparison of diagnostic accuracy of ultrasound and mammography in detecting breast cancer in radiographically dense breasts. Cureus. 2025;17(9). doi:10.7759/cureus.92637
14. Quratulain S, Bibi S, Gul S, Khatoon S, Mujeeb Z, Akram S. The diagnostic accuracy of conventional breast ultrasound in diagnosing malignant breast lesions taking histopathology as gold standard. Pak J Health Sci. 2024;5(6):36–41. doi:10.54393/pjhs. v5i06.1657
15. Alotaibi SA, Gareeballah A, Sami Ur Rahman J, Faqihi WA, Ali S, Elbashir M, et al. Evaluation of the diagnostic accuracy of mammography and ultrasonography for breast cancer diagnosis using histopathology as the reference standard: a cross-sectional study. Ann Med Surg. 2026;88(1):159–66. doi:10.1097/MS9.0000000000004369

16. Iovănescu G, Bîrsăşteanu F, Borugă VM, Apostol A, Ştefănescu EH, Budu VA, et al. Clinical, ultrasound and histopathological correlation of clinically N0 neck nodes in patients with cancers of the pharynx and larynx. Rom J Morphol Embryol. 2020;61(2):433–9. doi:10.47162/RJME.61.2.12
17. Akinnibosun-Raji HO, Saidu SA, Mustapha Z, Ma'aji SM, Umar M, Kabir FU, et al. Correlation of sonographic findings and histopathological diagnoses in women presenting with breast masses. J West Afr Coll Surg. 2022;12(2):109–14. doi: 10.4103/jwas.jwas_84_22
18. Xiang W, Rao H, Zhou L. A meta-analysis of contrast-enhanced spectral mammography versus MRI in the diagnosis of breast cancer. Thorac Cancer. 2020;11(6):1423–32. doi:10.1111/1759-7714.13400
19. Nkya GZ, Omoyiola OZ, Betiku OA, Ng DL, Fernandes FC, Tacula NS, et al. Assessment of breast pathology reporting needs and development of tumor synoptic templates in sub-Saharan Africa. Arch Pathol Lab Med. 2025;149(4):340–6. doi:10.5858/arpa.2024-0101-OA
20. Eghtedari M, Chong A, Rakow-Penner R, Ojeda-Fournier H. Current status and future of BI-RADS in multimodality imaging, from the AJR special series on radiology reporting and data systems. AJR Am J Roentgenol. 2021;216(4):860–73. doi:10.2214/AJR.20.24894
21. Yoon JH, Kim MJ, Lee HS, Kim SH, Youk JH, Jeong SH, et al. Validation of the fifth edition BI-RADS ultrasound lexicon with comparison of fourth and fifth edition diagnostic performance using video clips. Ultrasonography. 2016;35(4):318–26. doi:10.14366/usg.16010
22. Nwadinigwe U, Itua I, Keyworth C, Haith-Cooper M. Interventions to improve breast screening uptake for Black African women in high income countries: a systematic review. J Racial Ethn Health Disparities. 2026 May 26. doi:10.1007/s40615-026-02999-9
23. Nwadinigwe U, Itua I, Haith-Cooper M, Keyworth C. Barriers and facilitators of breast cancer screening services amongst Black African women in high income countries: a systematic review and meta-aggregation. Discov Public Health. 2026;23(1):521. doi:10.1186/s12982-026-01561-6
24. Badu-Peprah A, Otoo OK, Amamoo M, Quarshie F, Adomako B. Breast imaging reporting and data system for sonography: positive and negative predictive values of sonographic features in Kumasi, Ghana. Transl Oncol. 2024; 45:101976. doi: 10.1016/j.tranon.2024.101976
25. Mubuuke AG, Nassanga R, Galukande M. Comparative accuracy of sonography, mammography and the BI-RADS characterization of breast masses among adult women at Mulago Hospital, Uganda. J Glob Health Rep. 2023;7: e2023013. doi:10.29392/001c.75139

Figures

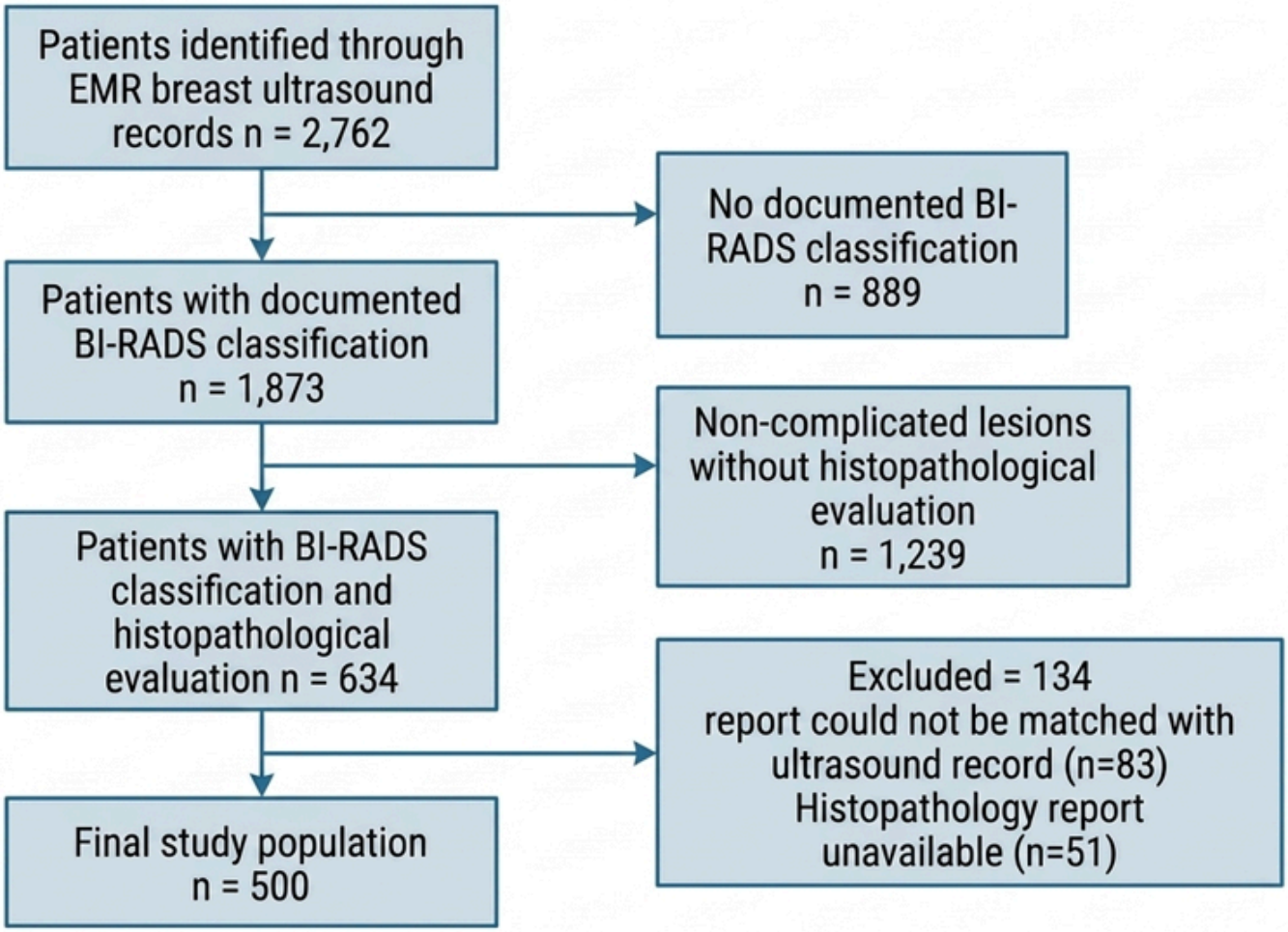


Figure 1

STARD flow diagram of patient selection for the diagnostic accuracy study of breast ultrasound compared with histopathology.

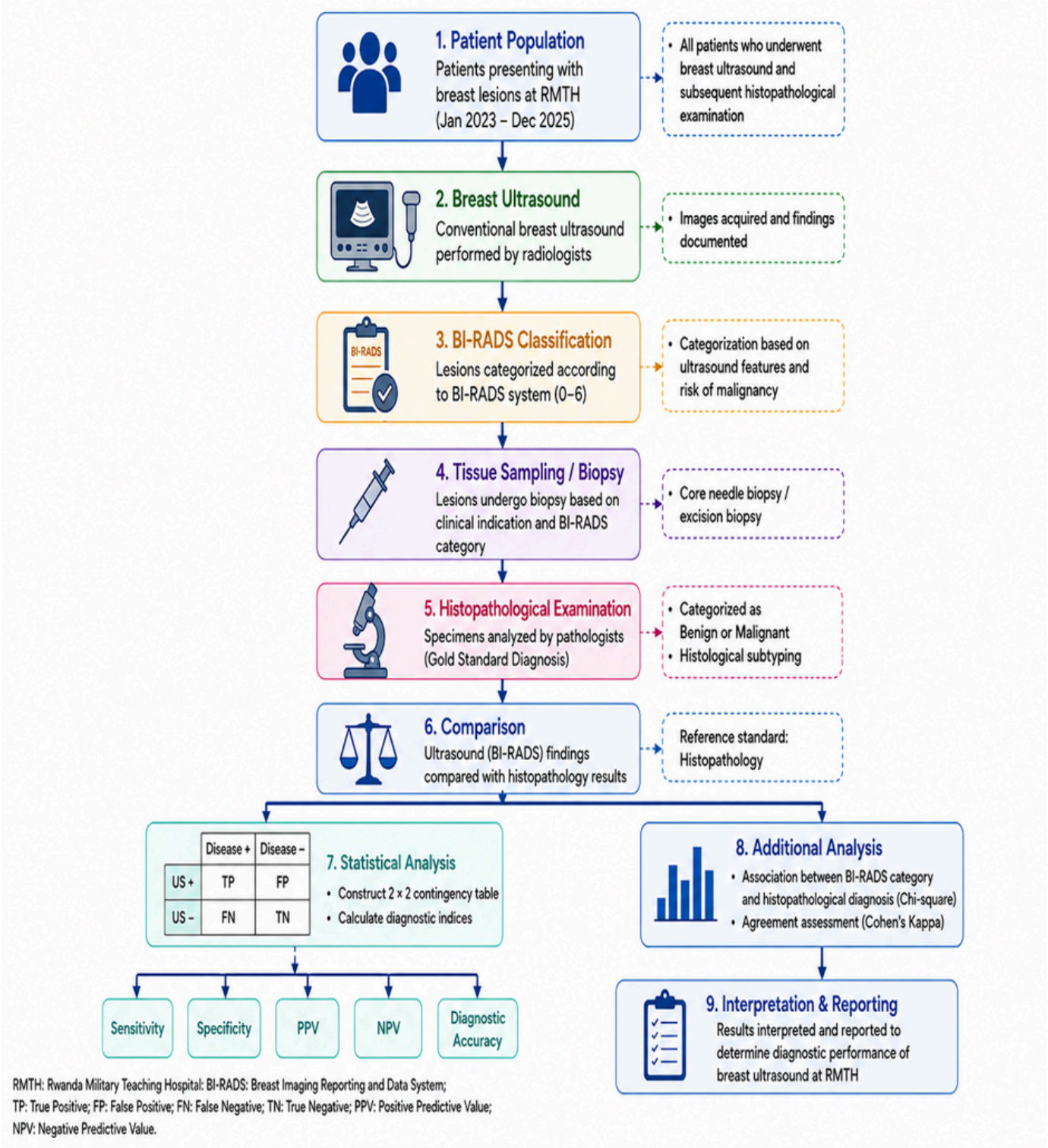


Figure 2

Study flow diagram from ultrasound examination to histopathological correlation.

Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.819	0.019	0.000	0.781	0.857

The test result variable(s): biradsforROC CURVE has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

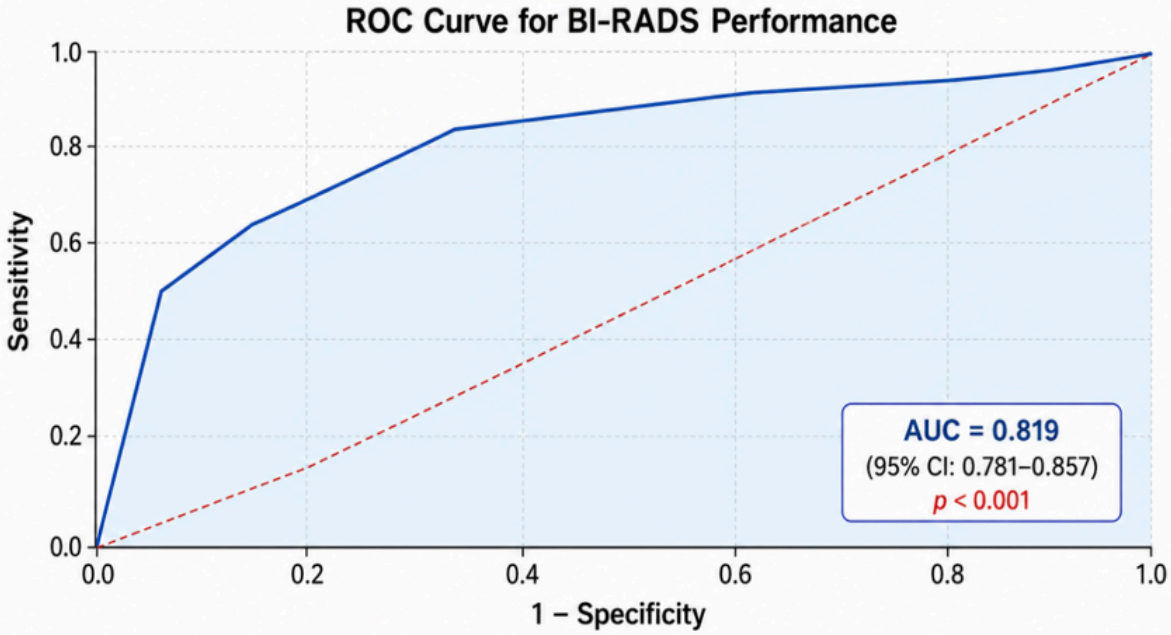


Figure 3

Receiver operating characteristic (ROC) curve demonstrating the performance of BI-RADS classification in discriminating malignant from benign breast lesions using histopathological diagnosis as the reference standard (AUC = 0.819, 95% CI 0.781–0.857; p < 0.001).