Accuracy of Recommendations and Histopathology Results with Breast Imaging Reporting and Data System (BI-RADS) – A Retrospective Study

Amir Yousaf[1], Naushina S. Zaidi[2], Fareeha Rauf[2], Raheel Ahmed[2], Madeeha R. Malik[3], Shawaz Abbas[4]

ABSTRACT

Objective: To evaluate the diagnostic accuracy of radiology team recommendations compared with histopathology outcomes using Breast Imaging Reporting and Data System (BI-RADS) categories in breast ultrasound (US).

Study Design: Retrospective cross-sectional study.

Place and Duration: Radiology Department, Mian Meer Hospital, Lahore, in collaboration with Mayo Hospital, Lahore, Pathology Department, from 02 January 2021 to 03 June 2022.

Methodology: Female patients referred from the Well Women Clinic (WWC) (N=361), which provides breast health services to the local community, and who underwent breast ultrasound performed by consultant radiologists, were included. Patients had USS scans by non-consultants. Male and female patients referred from sources other than the Well Women Clinic were excluded. BI-RADS categories 1–5 were assigned and compared with American College of Radiology (ACR) BI-RADS guidelines and with histopathology reports where available. Category 0 was excluded due to incomplete data. SPSS v29 was used for statistical analysis.

Results: Out of 48,432 ultrasound referrals, 539 were from WWC; 361 met the inclusion criteria. Mean age was 36.23 ± 13.18 years. BI-RADS distribution: Category 1 = 56.5%, Categories 2–3 = 19.1%, Categories 4–5 = 24.4%. Agreement with ACR BI-RADS guideline recommendations was significant (χ^2=475.3, p<0.0001), with accuracies of 100% for discharge, 97% for follow-up, and 97% for referral. Histopathology confirmed malignancy in 41 patients (11.4%). Sensitivity and specificity for malignancy detection were 76% and 95% respectively, with overall diagnostic accuracy of 93%.

Conclusion: BI-RADS-based breast ultrasound by consultant radiologists demonstrated high accuracy against histopathology and guideline recommendations and supporting its role as a reliable first-line diagnostic tool.

Keywords: Breast Imaging Reporting and Data System, Breast ultrasound, Histopathology correlation, Benign breast disease, Breast cancer diagnosis.

How to Cite This:

Yousaf A, Zaidi NS, Rauf F, Ahmed R, Malik MR, Abbas S. Accuracy of Recommendations and Histopathology Results with Breast Imaging Reporting and Data System (BI-RADS) – A Retrospective Study. Isra Med J. 2025; 17(1):23-27. DOI: https://doi.org/10.55282/imj.oa1425

INTRODUCTION

The radiological diagnoses of visceral pathologies started with the invention ultrasound scan (US). The first person to differentiate features of breast mass using US as benign and malignant in 1951 was Wild and Neal.[1] Later on, the use of US scan to detect sinister breast pathology increased with the improvement of its resolution and quality. Breast cancer is the most common malignancy in women worldwide, with increasing incidence in Pakistan.[2] Early and accurate detection significantly improves prognosis. Ultrasound (US) is widely available, especially in younger women with dense breasts and in low-resource settings where mammography is limited.

The American College of Radiology introduced the Breast Imaging Reporting and Data System (BI-RADS).[3] to standardise breast imaging interpretation and management. Initially developed for mammography.[4] BI-RADS has since been adapted for ultrasound, providing a structured lexicon and consistent reporting.

A dictionary of terms for describing lesions seen in the US is provided in the BI-RADS lexicon.[5]. The standardized assessment categories used to describe US scan findings are "0" for "need additional age-related evaluation", "1" for "negative", "2" for "beneficial finding", "3" for "probably benign finding", "4" for "suspicious abnormality", and "5" for "highly suspicious of a malignancy."

Although BI-RADS improves reproducibility, diagnostic accuracy may still vary depending on radiologist experience and institutional practices.[6] One of the few studies conducted in Pakistan has evaluated BI-RADS performance and concluded

1. Consultant Surgeon, Mian Meer Hospital Lahore
2. Consultant Radiologist, Mian Meer Hospital Lahore
3. Consultant Radiologist, District Hospital Nankana Sahib, Lahore
4. Medical Officer, Mian Meer Hospital Lahore

Correspondence:
Amir Yousaf
Consultant Surgeon, Mian Meer Hospital, Lahore.
E-mail: yousafamir@hotmail.com

Received for Publication: August 18, 2025
1st Revision of Manuscript: Sep 20, 2025
2nd Revision of Manuscript: Nov 30, 2025
Accepted for Publication: Dec 1. 2025

that sub-categorisation of BI-RADS category 4 should be practised routinely, as it is associated with a risk of malignancy. The main objective of this study was to evaluate the accuracy of BI-RADS-based recommendations by consultant radiologists at Mian Meer Hospital, Lahore, compared with the histopathology results.

METHODOLOGY

This retrospective cross-sectional study was conducted in the Radiology Department of Mian Meer Hospital, Lahore, with corresponding histopathology records obtained from Mayo Hospital, a tertiary care centre. It was carried out from 2nd January 2021 to 3rd June 2022, after getting IRB approval i.e. letter no 4124-29/GMMH dated 4th August 2022.

A consecutive sampling technique was used for study proceedings. Female patients referred from the WWC for breast ultrasound (by consultant radiologists), with available histopathology reports for biopsy during the study period, were included. Patients referred from other departments of the hospital, male patients, ultrasounds performed by non-consultant radiologists and BI-RADS category 0 cases with incomplete follow-up were excluded. To assess the diagnostic accuracy of the reported BI-RADS categories against final tissue diagnosis, histopathology was considered the gold standard for benign or malignant disease. For that, BI-RADS 1 to 3 were interpreted as negative for malignancy, while BI-RADS 4 and 5 were treated positive. For each recommendation group, true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) were identified. Using these values, sensitivity, specificity, likelihood ratios and overall diagnostic accuracy were calculated to determine how well BI-RADS-based recommendations predicted true disease status.

Data collection was carried out by obtaining access from the record department of Mian Meer Hospital, upon request by the surgical medical officer. While the pathology department of Mayo Hospital, Lahore, provided corresponding biopsy histopathology data upon request by the medical superintendent. The demographic details, BI-RADS category (1 to 5), breast side and quadrant, and histopathology findings from FNAC or excisional biopsy were documented.

To minimise inter-observer variation, only consultant radiologist reports were analysed. Consecutive sampling further reduced selection bias by ensuring that all eligible cases within the defined timeframe were included. Additionally, matching radiology data with independently reported histopathology results helped prevent diagnostic incorporation bias, as the pathologists were unaware of the study objectives. These combined methods strengthened the internal validity of the study.

Data Analysis: The statistical analysis was performed by using Microsoft Excel 2016 for age categorisation, consultant radiologist codes and distribution of breast pathologies according to BI-RADS categories, while SPSS v29 was used for descriptive statistics. Pearson chi-square test with a p-value less than 0.05 considered statistically significant, assessed associations between BI-RADS categories, recommendations and histopathology outcomes.

RESULTS

A total of 48,432 (100%) patients were referred to the radiology department from all the departments of the hospital. Out of the total number of patients, 539 (1.11%) patients were referred from WWC for a breast US scan. After exclusion of the remaining 361 patients, with a mean age of 36.23 ± 13.18 years, the data were divided into two age groups: group 'A' (patients aged <35 years) and group 'B' (patients aged>35 years). Reviewing the data through BI-RADS categories (Table I), it was observed that normal scans predominated (56.5%), indicating that the majority had no suspicious findings. Fibro adenomas were mostly assigned BI-RADS 2–3 (10.5%), while irregular lesions were concentrated in BI-RADS 3–4 (8.3%). Rare findings like dilated ducts (0.6%) and "lumps & bumps" (24.1%) were distributed across multiple categories.

Table I: Breast pathologies reported as BI-RADS categories by the radiology team (N=361)

Pathology	BI-RADS Categories				
	1 n(%)	2 n(%)	3 n(%)	4 n(%)	5 n(%)
Normal	204 (56.51)	0 (0)	0 (0)	0 (0)	0 (0)
Fibro adenoma	0 (0)	31 (8.59)	7 (1.94)	0 (0)	0 (0)
Irregular lesion	0 (0)	1 (0.28)	19 (5.26)	10 (2.77)	0 (0)
Dilated ducts	0 (0)	0 (0)	1 (0.28)	0 (0)	1 (0.28)
Lumps and Bumps	2 (0.55)	55 (15.24)	27 (7.48)	3 (0.83)	0 (0)

Note: BI-RADS = Breast Imaging Reporting and Data System.

It was observed that the total number of left breasts scanned was 205 (56.79%) as compared to 152 (42.11%) for right breasts. The younger women showed a higher proportion of benign lesions (28.7%), with no malignant cases recorded (Table II). In contrast, older women presented a significant proportion of malignancies (23.3%), along with a reduced benign rate (12.5%). Group 'A' showed only benign pathologies (45.7%), while Group 'B' had 41 malignant cases (23.3%). The chi-square analysis (χ^2=55.09, p<0.0001) confirmed statistical significance. This shows that malignancies were strongly associated with older age, emphasising age as a risk factor for breast cancer.

Table II: Breast pathologies reported in age groups (N=361)

Pathology	Group A n(%)	Group B n(%)	χ^2	p – value
Benign	53 (28.65)	22 (12.50)	55.09	<0.0001
Discharged	110 (59.46)	94 (53.41)		
Malignant	0 (0)	41 (23.30)		
Non – Compliant	22 (11.89)	19 (10.80)		
Overall	185 (100)	176 (100)		

Note: χ^2 = Pearson chi-square test.

DOI: https://doi.org/10.55282/imj.oa1425

Table III. BI-RADS guideline recommendations and radiologists' reports (N=361)

Guideline Recommendations	Reported Recommendations			χ^2	p – value
	Discharge n(%)	Follow-up n(%)	Referred n(%)		
Normal	206 (100.00)	0 (0)	0 (0)	475.31	<0.0001
Follow-up	0 (0)	118 (100)	23 (62.16)		
Tissue Diagnosis	0 (0)	0 (0)	14 (37.84)		
Overall	206 (100)	118 (100)	37 (100)		

Table IV. Correlation between histopathology results and radiologists' recommendations (N=361)

Tissue Diagnosis	Reported Recommendations			χ2	p – value
	Discharged. n(%)	Follow-up. n(%)	Referred. n(%)		
Malignant	**0 (0)**	**16 (13.56)**	**25 (67.57)**	451.05	<0.0001
Benign	2 (0.97)	65 (55.09)	8 (21.62)		
Non – Compliant	0 (0)	37 (31.36)	4 (10.81)		
Discharged	204 (99.03)	0 (0)	0 (0)		
Overall	206 (100)	118 (100)	37 (100)		

To analyse the BI-RADS categories reported by the consultant radiologist, the data were tabulated to compare with BI-RADS recommended guidelines (Table 3), which evaluates concordance between radiologists' recommendations and BI-RADS standards. Overall, there was a strong correlation between radiologists' recommendations and ACR BI-RADS guidelines, supporting standardised and reliable reporting (χ^2=475.3, p<0.0001).

The tissue histopathology results were marked as benign and malignant and compared with the recommendations or actions taken by the radiology team, which were labelled as 'referred' patients for highly suspected breast lesions as BI-RADS categories 4 and 5. Categories 2 and 3 were marked as 'follow-up' patients, and category 1 as 'discharged'. Malignancies were predominantly in referred cases (67.6%), while benign lesions were mostly in follow-up (55.1%). Discharged patients almost exclusively had normal histopathology (99%) (Table 4). Statistical analysis confirmed significance (χ^2=451.05, p<0.0001).

To evaluate the diagnostic accuracy indicators, the 'Discharge group' was identified as benign (TN = 206), and no false positives or false negatives were recorded (FP = 0, FN = 0). The 'Follow-up group' was flagged as (TP = 118), while 206 benign lesions were categorised (TN = 206). A total of 23 cases were false positives (FP = 23), with no false negatives (FN = 0). 'Referred for Tissue Diagnosis group', histopathology confirmed 14 true malignant cases (TP = 14) and 23 benign cases (TN = 23). Thirteen cases classified as requiring tissue diagnosis were ultimately benign, representing false negatives (FN = 13), while no false positives were observed (FP = 0).

The resulting diagnostic performance parameters were calculated using standard formulas as follows;

- Sensitivity:1.00 for Discharge and Follow-up groups, and 0.52 for the Referred group.
- Specificity:1.00 for Discharge, 0.95 for Follow-up and 1.00 for the Referred group.
- Diagnostic Accuracy:1.00 for Discharge, 0.97 for Follow-up and 0.97 for the Referred group.

Because both the Discharge and Referred groups had perfect specificity, likelihood ratios tended toward infinity, indicating strong discriminatory ability, while the Follow-up group demonstrated moderately reduced LR values due to the presence of false positives.

Finally, a Receiver Operating Characteristic (ROC) curve was created using BI-RADS category as the test variable and histopathology as the gold standard. The area under the ROC curve (AUC) quantified the overall ability of BI-RADS recommendations to differentiate benign from malignant lesions. A high AUC value provided pictorial and statistical evidence supporting the diagnostic accuracy of the recommended BI-RADS reports by the consultants in this patient cohort.

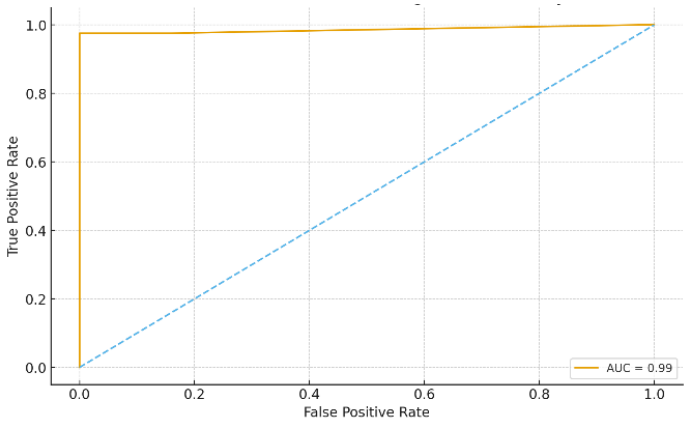

Fig 1: ROC Curve for BI-RADS Diagnostic Accuracy

DISCUSSION

The BI-RADS jargon was first evolved in 1993 for reporting mammography.[8] Breast sonography is now well established as a useful and improving technique, despite the fact that mammography is acknowledged as the most effective method for screening for breast cancer. A number of studies have shown that sonographic appearance can help distinguish between benign and malignant solid breast masses. Our study demonstrates that BI-RADS–based ultrasound reporting by

DOI: https://doi.org/10.55282/imj.oa1425

consultant radiologists achieves high diagnostic accuracy, with a sensitivity of 76% and a specificity of 95%, which is comparable to international standards. Similar findings have been documented in global and regional studies, reinforcing BI-RADS as a reliable framework for breast cancer diagnosis.[9]

The findings in our experience showed that the reported management recommendations were very similar to those suggested by the guideline recommendations. In a few cases, the radiologists were very precise by assessing the lesions with subclasses 'A', 'B' and 'C' in BI-RADS categories 4 and 5. Moreover, our radiology team was taking extra precautions in the care of their patients by referring them for FNAC or to surgeons for excisional biopsies. Our malignancy rate of 11.4% aligns with regional screening data. Importantly, benign lesions were appropriately categorised, minimising unnecessary biopsies. Non-compliance in biopsy referrals highlights a limitation in patient follow-up. The observed aligns with reported prevalence rates in South Asia, underscoring the importance of accurate triaging in younger women and in low-resource environments.[7]

Recent research emphasises that adjunct imaging methods can improve diagnostic performance. Superb Microvascular Imaging has demonstrated enhanced differentiation of benign versus malignant lesions.[10], while sono elastography provides valuable information on lesion stiffness.[11] Radiomics and artificial intelligence are increasingly being integrated into BI-RADS scoring, with deep learning models showing AUC values above 0.90 in distinguishing BI-RADS 4 lesions.[12-14] Importantly, AI-supported diagnostic systems have improved inter-observer agreement among radiologists of different expertise levels.[15-16]

A 2021 Iranian study reported average diagnostic accuracy of 90.15% for BI-RADS 3-5 ultrasound.[17] The strengths of our study include restriction to consultant radiologists, minimising observer variability. Limitations include the single-centre retrospective design, modest sample size, and some non-compliance among biopsy-referred patients. Future multicentre studies incorporating AI, elastography, and radiomics would further validate these findings across diverse populations.

ACKNOWLEDGEMENTS: Dr Muhammad Akhtar (Medical Superintendent of Mian Meer Hospital, Lahore), who assisted in obtaining the histopathology results from the pathology department of the tertiary hospital, Mayo Hospital, Lahore.

CONCLUSION

Breast ultrasound using BI-RADS lexicon by consultant radiologists showed high diagnostic accuracy compared to histopathology. It is an effective first-line investigation in both younger and older women, particularly where access to mammography is limited.

AUTHOR'S CONTRIBUTION

Yousaf A: Conceived the idea, Designed Research Methodology, manuscript writing, Proofreading & Final Approval
Zaidi NS: Designed Research Methodology
Rauf F: Manuscript Writing
Ahmed R: Data Analysis
Malik MR: Data Analysis
Abbas S: Data Collection

Disclaimer: None.
Conflict of Interest: None.
Source of Funding: None.

REFERENCES

1. Merz E, Evans DH, Dong Y, Jenssen C, Dietrich CF. History of ultrasound in obstetrics and gynaecology from 1971 to 2021 on occasion of the 50 years anniversary of EFSUMB. Med Ultrason. 2023; 25(2):175-188.
2. Siddiqui R, Mehmood MH, Khan NA. An overview of breast cancer in Pakistan. Discov Med. 2024; 1(1):82.
3. American College of Radiology. BI-RADS Atlas. 5th ed. Reston (VA): ACR; 2017.
4. Ghaemian N, Tehrani NHG, Nabahati M. Accuracy of mammography and ultrasonography BI-RADS in breast cancer. Caspian J Intern Med. 2021; 12(2):S350-S356.
5. 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):318-326.
6. Xu Z, Zhong S, Gao Y, Huo J, Xu W, Huang W, et al. Optimizing breast lesions diagnosis and decision-making with a deep learning fusion model integrating ultrasound and mammography: a dual-center retrospective study. Breast Cancer Res. 2025; 27(1):80.
7. Zahra SB, Majeed AI, Malik N. Positive predictive rate of BI-RADS category 4 lesions for breast cancer taking histopathology as gold standard. Pak Armed Forces Med J. 2025; 75(Suppl-2): S202-S205.
8. Urban L. BI-RADS® Classification. In: Novita G, Frasson A, Millen E, Zerwes F, Cavalcante F. (eds) Breast Diseases. Springer, Cham.2019;1-7. DOI:https://doI.org/10.1007/978-3-030-13636-9_1
9. Ghaemian N, Tehrani NHG, Nabahati M. Accuracy of mammography and ultrasonography and their BI-RADS in detection of breast malignancy. Caspian J Intern Med. 2021; 12(4):573.
10. Feng J, Lu J, Jin C, Chen Y, Chen S, Guo G, et al. Diagnostic value of superb microvascular imaging in differentiating benign and malignant breast tumors: a systematic review and meta-analysis. Diagnostics. 2022; 12(11):2648.
11. Selim YAY, Sabit H, Arneth B, Shaaban MA. Diagnostic accuracy of sonoelastography for breast lesions: a meta-analysis comparing strain and shear wave elastography. J Imaging. 2025; 11(7):221.
12. Dan Q, Xu Z, Burrows H, Bissram J, Stringer JS, Li Y. Diagnostic performance of deep learning in ultrasound diagnosis of breast cancer: a systematic review. NPJ Precis Oncol. 2024; 8(1):21.

13. Yang Y, Liao T, Lin X, Ouyang R, Cao Z, Hu J, et al. Reducing unnecessary biopsies of BI-RADS 4 lesions based on a deep learning model for mammography. Front Oncol. 2025; 15:1543553.
14. Wang Y, Tang L, Chen P, Chen M. The role of a deep learning based computer-aided diagnosis system and elastography in reducing unnecessary breast lesion biopsies. Clin Breast Cancer. 2023; 23(3):e112-e121.
15. Çelebi F, Tuncer O, Oral M, Duymaz T, Orhan T, Ertaş G. Artificial intelligence in diagnostic breast ultrasound: decision support among radiologists with various levels of expertise. Eur J Breast Health. 2025; 21(1):33-40.
16. Tsuyuzaki S, Fujioka T, Yamaga E, Katsuta L, Mori M, Yashima Y, et al. Evaluation of AI diagnostic systems for breast ultrasound: comparative analysis with radiologists and the effect of AI assistance. Jpn J Radiol. 2025. 43(10):1643-1651. DOI: 10.1007/s11604-025-01809-2
17. Ghaemian N, Tehrani NHG, Nabahati M. Accuracy of mammography and ultrasonography and their BI-RADS in detection of breast malignancy. Caspian J Intern Med. 2021; 12(4):573.

 DOI: https://doi.org/10.55282/imj.oa1425