Original Article

Assessment of BI-RADS Reliability in Sonomammographic Diagnosis of Breast Masses Confirmed by Postoperative Histopathology: A Multicenter Observational Study

Hazem Ghanem[1], Ibrahim Abdelrafeh[1], Mohammad Moneir[1] and Mohammad Abobakr[2]

[1]Department of General Surgery, [2]Department of Diagnostic and Intervention Radiology, Armed Forces College of Medicine, Cairo, Egypt

ABSTRACT

Background: Breast lumps are a common presentation of both benign and malignant breast lesions. Triple assessment improves diagnostic accuracy, and the Breast Imaging Reporting and Data System (BI-RADS), based on sonomammography, provides standardized reporting and malignancy risk stratification. This study assesses the diagnostic accuracy of BI-RADS in breast masses.

Methods: This observational study included 77 surgical patients with breast masses. Sonomammography was performed using standard high-frequency linear transducers. Multiple experienced radiologists independently assigned BI-RADS categories per ACR guidelines, blinded to clinical and histopathological findings. Discrepancies were resolved by consensus. Histopathology of surgical specimens served as the reference standard.

Results: In benign lesions, BI-RADS categories were: 1 case as 5, 12 as 4b, 3 as 4c, 8 as 3, and 14 as 4a. Among malignant lesions, 0 were 3, 2 were 4a, 3 were 4b, 11 were 4c, and 23 were 5. Malignancy rates were 0% in category 3, 12.5% in 4a, 20% in 4b, 78.6% in 4c, and 95.8% in 5. Sensitivity was 87.2%, specificity 89.5%, positive predictive value 89.5%, negative predictive value 87.2%, and overall diagnostic accuracy 88.3%.

Conclusion: The study confirms a strong relationship between higher BI-RADS categories and malignancy, reinforcing its diagnostic utility in surgical cohorts. The exclusive use of histopathologically confirmed cases adds precision, confirming BI-RADS 3 as reliably benign. These findings support BI-RADS as a valuable tool for surgical decision-making and highlight the need for larger, multi-institutional studies.

Key Words: BI-RADS, breast cancer, postoperative-histopathology.

Received: 21 May 2025, **Accepted:** 26 May 2025

Corresponding Author: Hazem Gamal Ghanem, Department of General Surgery, Armed Forces College of Medicine, Cairo, Egypt, **Tel.:** +2 015 5646 0060, **E-mail**: Hazemgamalghanem@gmail.com

ISSN: 2812-5509, 2025

INTRODUCTION

A lump can form as a result of any type of breast tumour, benign or malignant. Breast cancer accounts for more than 25% of all female cancers worldwide, impacting more women than any other type of cancer[1]. Pathological diagnosis, clinical examination, and radiological imaging (mammography, ultrasonography) can all help to improve the final diagnosis' accuracy. However, not all malignant breast masses become benign, and not all benign breast lumps develop into cancer. They can be used for both screening and diagnosis[2].

Breast lesions are now much easier to identify because to advances in imaging technology. Early identification, medication, and a favourable prognosis all contribute to higher survival rates for breast cancer patients. Ultrasonography and mammography are two noninvasive, widely available, and fairly cost radiological treatments that aid in the diagnosis process[3]. Any woman over the age of 40 who develops a lump in her breast should get a mammogram to be sure it is not cancer[4].

The Breast-Imaging and Reporting Data System (BI-RADS) is commonly used for reporting by breast imaging modalities such as mammography, MRI, and ultrasound. The reporting can be thought of as assigning a BI-RADS category score and then informing appropriate management. BI-RADS for mammography and ultrasonography includes the following data: The evaluation process utilises the following categories: 0 denotes an incomplete assessment, 1 suggests a negative finding, 2 indicates benign results, 3 indicates likely benign findings, 4 indicates suspected abnormalities, and 5 indicates a high risk of malignancy[5].

Group 3 is the BI-RADS group with the lowest cancer risk (less than 2%). BI-RADS class 4 predicts approximately 30% of breast cancer diagnoses, while class 5 predicts more than 95% of all cancer cases. The BI-RADS 4a, 4b, and 4c subcategories are used to further stratify the cancer risk in category 4. The American College of Radiology (ACR) recommends multiple therapy options for each available category. We will use BI-RADS to evaluate categories 1 and 2. Patients in category 3 BI-RADS should have a

DOI: 10.21608/ARCMED.2025.387688.1129

brief interval follow-up every six months, but patients in categories 4 and 5 should have tissue diagnostics[6].

The concern with Categories 3 and 4 is that they strike a balance between the need for thorough monitoring and the risk of missing early-stage cancers. To achieve the goals of reducing unnecessary therapies and ensuring timely detection and management of breast cancer, surgeons must carefully analyse clinical factors, imaging findings, and patient history when classifying lesions into these groups.

In Egypt, there is a scarcity of study on the reliability of BiRADs. In this study, we compared the final diagnosis of the histopathoiogical diagnosis to the results of our institution's radiological evaluation. The inquiry also revealed information on the test's specificity, accuracy, and sensitivity levels. This study aims to demonstrate that the BI-RADS categories are highly predictive and to analyse how they influence surgical intervention and treatment outcomes.

MATERIAL AND METHODS

This was an observational analytical follow-up study. Female patients were selected and subjected to the following:

Preoperative

History taking and clinical evaluation of the women were conducted, followed by analysis of their sonomammogram findings with respect to BI-RADS. The evaluation process utilized the following categories: 0 denoted an incomplete assessment, 1 suggested a negative finding, 2 indicated benign results, 3 indicated likely benign findings, 4 indicated suspected abnormalities, and 5 indicated a high risk of malignancy.

Operative

The breast mass samples were preserved by applying suitable techniques for their removal.

Postoperative

Histopathological examination was performed to assess the diagnostic accuracy of BI-RADS in identifying breast masses, using histopathology as the gold standard.

Ethical considerations

- The Ethical Review Committee of the Armed Forces College of Medicine reviewed the amended research proposal and gave its approval.
- Each patient was informed of the purpose and nature of the research prior to their involvement, and data confidentiality was maintained at all times.
- Before enrolling each participant, an informed written consent will be obtainedfrom all participants before enrollment. The study design conformed to the requirementsof Revised Helsinki Declaration of biomedical ethics.
- **Confidentiality of data:** Patients' data will be dealt with in complete confidentiality, and no one has right to read their medical information except the investigators in this study. After the research is complete, they will be informed regarding their results and also further information regarding their health status. Individual confidentiality will be maintained in all published and written data resulting from the study
- **Right to refuse or withdraw:** Any participant doesn`t have to take part in this research if he doesn't want. They may also stop participating at any time without any affection to the medical care provided.

Research design and setting

Study design: Observational analytical follow up study.

Study setting: Participants were recruited from the general surgery outpatient clinic of [AFCM hospital, Ghamra military hospital and Maadi military hospital medical records; health registers (including; history taking; clinical examination and radiological evaluation)

Participants

Female patients presented with breast mass and admitted to the general surgery department in AFCM hospital, Ghamra military hospital and Maadi military hospital

Inclusion criteria

Females aged › 35 with breast mass submitted to surgery.

Exclusion criteria

- Recurrent cases after previous surgery
- Residual tumor after surgery

Data collection tools

- History and clinical examination of women with the inclusion criteria
- Sonomammography was performed using standard high-frequency linear transducers by multiple experienced radiologists who independently reviewed the images blinded to clinical and histopathological outcomes
- A variety of operations are performed including mastectomy, breast-conserving surgery (lumpectomy),

- Postoperative histopathology reports

Procedures

All patients were subjected to the following:

- Full history taking, proper clinical examination
- Radiological evaluation: sonomammography was performed using standard high-frequency linear transducers by multiple experienced radiologists who independently reviewed the images blinded to clinical and histopathological outcomes
- Laboratory Investigations include routine and necessary preoperative investigations
- Breast surgery according to the patients case
- Obtain post-operative histopathology reports

Sampling and Sample size

Sample type: Selection of non-probabilistic convenience samples from military hospitals in Maadi, Ghamra, and AFCM

Sample size: the total required sample to be enrolled in the study is 77 female patients

Statistical analysis

Pre-coded data were processed and statistically analysed using the Statistical Package for the Social Sciences (SPSS), version 21. Mean, standard deviation, median, and interquartile range (IQR) were used to describe quantitative data, while number and percent were used for qualitative data. When comparing qualitative variables between two groups, the chi-square test was used; for comparing quantitative variables, the independent t-test was applied. When necessary, additional statistical tests were used. A statistically significant result was defined as a *P value* below 0.05.

RESULTS

(Table 1) summarizes the demographic and clinical characteristics of the 77 female patients included in this study. The mean age of the cohort was 54.8 ± 9.7 years, with patients diagnosed with malignant breast masses being significantly older than those with benign lesions (58.9 ± 9.1 years vs. 50.7 ± 8.5 years; p = 0.001). The mean body mass index (BMI) was 27.3 ± 3.8 kg/m², with no statistically significant difference observed between benign and malignant groups (p = 0.11).

Cardiovascular comorbidities were more prevalent in the malignant group, including hypertension (51.3% vs. 26.3%; p = 0.02), diabetes mellitus (38.5% vs. 18.4%; p = 0.04), and smoking (25.6% vs. 7.9%; p = 0.03). All patients presented with a palpable breast mass. Pain was more frequently reported among patients with benign lesions compared to those with malignancies (63.2% vs. 25.6%; p = 0.001). Conversely, clinical signs indicative of malignancy such as bloody nipple discharge (23.1% vs. 5.3%; p = 0.01) and ulceration (12.8% vs. 2.6%; p = 0.04) were significantly more common in the malignant group.

Furthermore, the interval from symptom onset to presentation was significantly longer in patients with malignant tumors (6.2 ± 2.7 months) compared to those with benign lesions (4.4 ± 2.0 months; p = 0.003). These findings delineate clear demographic and clinical distinctions between benign and malignant breast masses in this surgical cohort, underscoring the relevance of these parameters in preoperative evaluation.

Regarding the relationship between BI-RADS classification and postoperative histopathology.

For the benign lesions BI-RADS 3 were 8 patients, 4a were 14 patients, 4b were 12 patients, 4c were 3 patients, and BI-RADS 5 was one lesion benign.

While, BI-RADS 3 there was no malignant lesions, BI-RADS 4a were 2 malignant lesions, 4b were 3 malignant lesions, 4c were 11 malignant lesions, while BI-RADS 5 were 23 patients, (Table 2) (Figure 1). There was statistically significant difference between benign and malignant tumors regarding BI-RADS classification, p=0.001.

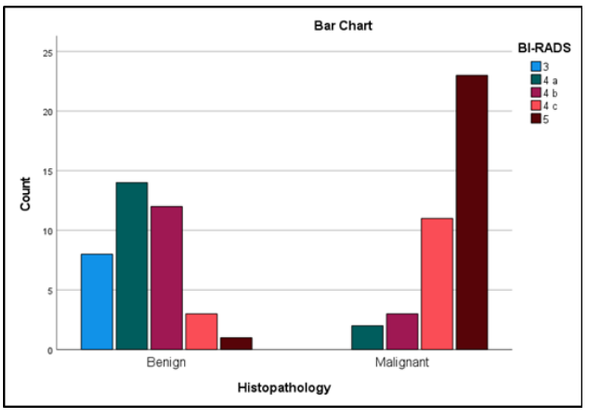

Fig. 1: Benign and malignant lesions to BI-RADS classification

The following table showed BI-RADS classification in correlation to diagnosis, (Figure 2) (Table 3). There was statistically significant difference between different histopathological subtypes regarding BI-RADS classification, p=0.001.

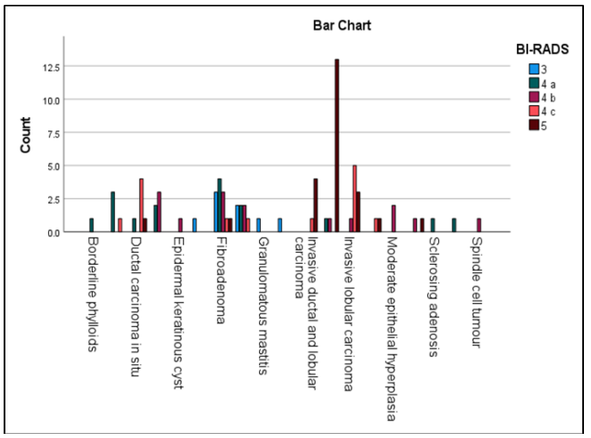

Fig. 2: correlation of pathological subtypes to BI-RADS classification

Sensitivity of BI-RADS Compared to Histopathological Findings

Based on histopathological correlation, the BI-RADS classification method demonstrated varying sensitivity levels for benign and malignant tumours. No malignant cases were found in BI-RADS category 3, indicating a high negative predictive value for benign lesions. The sensitivity of BI-RADS 4a was 12.5% for malignant lesions and 87.5% for benign ones.

BI-RADS 4b showed a sensitivity of 20% for malignant lesions and 80% for benign ones. BI-RADS 4c demonstrated a notable increase in sensitivity for malignancy (78.6%) compared to lower categories, with 21.4% representing benign cases. As expected, BI-RADS 5 exhibited a sensitivity of 95.8% for malignant tumours and 4.2% for benign lesions.

The BI-RADS system's diagnostic performance reflected these trends. When categories 4c and 5 were considered indicative of high suspicion for malignancy, the method achieved a sensitivity of 87.2% and a specificity of 89.5%. The positive predictive value (PPV) was 89.5%, the negative predictive value (NPV) was 87.2%, and the overall diagnostic accuracy reached 88.3%.

Table 1: Demographic and Clinical Characteristics

Variable	Total (n = 77)	Benign (n = 38)	Malignant (n = 39)	*p-value*
Age (years)				
Mean ± SD	54.8 ± 9.7	50.7 ± 8.5	58.9 ± 9.1	0.001*
Median (range)	53 (38–78)	49 (38–68)	59 (42–78)	
Body Mass Index (kg/m²)				
Mean ± SD	27.3 ± 3.8	26.7 ± 3.4	27.9 ± 4.0	0.11
Cardiovascular Risk Factors				
Hypertension, n (%)	30 (39.0%)	10 (26.3%)	20 (51.3%)	0.02*
Diabetes mellitus, n (%)	22 (28.6%)	7 (18.4%)	15 (38.5%)	0.04*
Smoker, n (%)	13 (16.9%)	3 (7.9%)	10 (25.6%)	0.03*
Clinical Presentation				
Palpable mass, n (%)	77 (100%)	38 (100%)	39 (100%)	–
Pain, n (%)	34 (44.2%)	24 (63.2%)	10 (25.6%)	0.001*
Bloody nipple discharge, n (%)	11 (14.3%)	2 (5.3%)	9 (23.1%)	0.01*
Ulceration, n (%)	6 (7.8%)	1 (2.6%)	5 (12.8%)	0.04*
Time Since Onset (months)				
Mean ± SD	5.3 ± 2.5	4.4 ± 2.0	6.2 ± 2.7	0.003*

Table 2: Cross-tabulation of BI-RADS Categories in relation to final histopathological diagnosis

Diagnosis	BI-RADS 3	4a	4b	4c	5	Total
Benign (n = 38)	8	13	11	3	3	38
Malignant (n = 39)	0	3	4	11	21	39
Total	8	16	15	14	24	77
% Malignant per BI-RADS	0.0%	18.8%	26.7%	78.6%	87.5%	–
% Benign per BI-RADS	100.0%	81.2%	73.3%	21.4%	12.5%	–

Table 3: Subtypes of diagnosis in relation to BI-RADS

Diagnosis	BI-RADS 3	BI-RADS 4a	BI-RADS 4b	BI-RADS 4c	BI-RADS 5	Total	% of Total (77 cases)
Borderline phyllodes	0	1	0	0	0	1	1.3%
Duct ectasia	0	3	0	1	0	4	5.2%
Ductal carcinoma in situ	0	1	0	4	1	6	7.8%
Ductal epithelial hyperplasia	0	2	3	0	0	5	6.5%
Epidermal keratinous cyst	0	0	1	0	0	1	1.3%
Fat necrosis	1	0	0	0	0	1	1.3%
Fibroadenoma	3	4	3	1	1	12	15.6%
Fibrocystic changes	2	2	2	1	0	7	9.1%
Granulomatous mastitis	1	0	0	0	0	1	1.3%
Intraductal papilloma	1	0	0	0	0	1	1.3%
Mixed invasive ductal and lobular carcinoma	0	0	0	1	4	5	6.5%
Invasive ductal carcinoma	0	1	1	0	13	15	19.5%
Invasive lobular carcinoma	0	0	1	5	3	9	11.7%
Metastatic carcinoma	0	0	0	1	1	2	2.6%
Moderate epithelial hyperplasia	0	0	2	0	0	2	2.6%
Mucinous carcinoma	0	0	1	0	1	2	2.6%
Sclerosing adenosis	0	1	0	0	0	1	1.3%
Sclerosing papilloma	0	1	0	0	0	1	1.3%
Spindle cell tumor	0	0	1	0	0	1	1.3%
Total	8	16	15	14	24	77	100%

DISCUSSION

The purpose of this study was to determine the reliability between preoperative BI-RADS (Breast Imaging Reporting and Data System) classifications and postoperative histopathological results in patients with suspicious breast lesions. This correlation is especially important since BI-RADS guides clinical judgements about the necessity for biopsy, surgical intervention, or follow-up imaging.

Our results showed an important correlation between higher BI-RADS categories and malignant histology. Specifically, cancer was identified in 0% of BI-RADS 3 lesions, 12.5% of BI-RADS 4a, 20% of BI-RADS 4b, 78.6% of BI-RADS 4c, and 95.8% of BI-RADS 5 lesions.

This pattern demonstrates the predictive aspect of the BI-RADS classification system, with increasing categories corresponding to higher levels of suspicion for cancer.

To further quantify the diagnostic value of BI-RADS, we evaluated its performance by grouping categories 4c and 5 as "high suspicion" (test-positive), and categories 3 to 4b as "low to moderate suspicion" (test-negative). Based on this stratification, BI-RADS achieved a sensitivity of 87.2% and a specificity of 89.5%. The positive predictive value (PPV) was 89.5%, the negative predictive value (NPV) was 87.2%, and the overall accuracy of the system in predicting malignancy was 88.3%.

These findings support BI-RADS as a reliable method for breast imaging risk stratification. Its reliability in ruling in and ruling out malignancy is supported by its high PPV and NPV, which helps surgeons choose the best course of action. While lower scores (BI-RADS 3 and 4a) were more commonly linked to benign pathology, enabling more conservative methods or short-term imaging follow-up, high scores (especially 4c and 5) were highly predictive of malignancy and required immediate biopsy or surgical action.

This study has limitations even though the diagnostic performance is encouraging. The results' generalisability may be impacted by the sample size's relative small size. Furthermore, this investigation did not assess interobserver variability in BI-RADS interpretation, a recognised problem in clinical practice that may affect classification consistency among various radiologists or institutions.

These results are in line with previous research regarding BI-RADS classification correlated with histological findings in breast tumors revealed strong association between cancer and higher BI-RADS categories, particularly BI-RADS 4c and 5, which was found in the research of 150 patients. The accuracy of BI-RADS in assessing cancer risk was highlighted by BI-RADS 5, which showed a malignancy rate of 96.43%. The authors emphasized the importance of BI-RADS 4 subcategories in the therapy of breast tumors and proposed more precise surgical decision-making[5].

The accuracy the Breast Imaging Reporting and Data System (BI-RADS) categories matched up with

histological results in breast cancer. It was a three-year, single-center retrospective study that looked at 316 breast specimens from 310 people. Category 3, 4, and 5 of the BI-RADS system were used to correlate imaging data with histological diagnosis in this investigation. Crucial findings include: Third BI-RADS even though this category is often associated with noncancerous findings, 2.8% of people in this study had malignancies identified.Additional study is needed in this area, since 26.6% of occurrences were malignant according to BI-RADS 4. The fact that 93.3% of cases were cancerous is further evidence of the high predictive value of this classification according to BI-RADS 5[7].

When deciding whether or not to perform a biopsy, surgeons might refer to the BI-RADS categorization. Because of this, patients with lower BI-RADS scores, who are at a higher risk of benign outcomes, may not undergo operations that are not absolutely essential.

The effectiveness of BI-RADS scoring systems in different healthcare settings (e.g., academic hospitals vs. community clinics) should be investigated in future research to identify the ways in which institutional factors impact diagnostic test outcomes.

Longitudinal studies that track the evolution of BI-RADS scores and how they relate to illness progression could further provide light on the system's predictive power. Increasing the sample size and include a broader variety of demographics would help reinforce the results[8].

Results demonstrate a strong correlation between BI-RADS scores and post-operative histology, particularly in the higher categories; this result lends credence to the BI-RADS system's ability to predict cancer. These findings support the ongoing use of BI-RADS grading in clinical settings to improve patient outcomes by demonstrating its therapeutic use in assessing breast lesions before surgery.

LIMITATIONS OF THE STUDY

- Interobserver variability: the accuracy and consistency of the results may be impacted by the variation among radiologists at various centres, and BI-RADS assessment is somewhat subjective.
- Lack of Long-Term Follow-Up: In order to evaluate the prognostic implications of BI-RADS categories, the study might not incorporate long-term patient outcomes like survival or recurrence.

CONCLUSION

This study emphasizes the therapeutic significance of BI-RADS classification in guiding the management of breast masses. By offering structured, evidence-based risk stratification, BI-RADS supports informed decisions on biopsy, surveillance, or surgery, ensuring timely and appropriate care—particularly in high-risk patients. The subcategorization within BI-RADS 4 further refines risk assessment, potentially reducing unnecessary invasive procedures in low-suspicion cases.

A novel insight from this study is the confirmation that BI-RADS 3 lesions were reliably benign within a surgically managed, histopathologically confirmed cohort, reinforcing the safety of conservative management in appropriately categorized cases. This adds precision to the clinical application of BI-RADS in surgical settings, where over-treatment is a concern.

Our findings affirm the diagnostic reliability of BI-RADS when interpreted collaboratively by surgeons and radiologists. However, its accuracy is influenced by radiologist experience and imaging quality. Interpretation variability across institutions remains a limitation. Additionally, the relatively small sample size in this multicenter study limits broad generalizability.

These findings should therefore be interpreted with caution. Future studies involving larger, diverse populations and standardized imaging protocols—with evaluation of interobserver consistency—are essential to validate and expand upon these results in routine clinical practice.

ABBREVIATIONS

AFCM: Armed Forces College of Medicine, **AJCC:** The American joint committee of cancer, **BI-RADs:** Breast Imaging-Reporting and Data System, **DCIS:** Ductal carcinoma in situ, **FSH:** Follicle-stimulating hormone, **IDC:** Invasive duct carcinoma, **ILC:** Invasive lobular carcinoma, **LCIS:** Lobular carcinoma in situ, **NST:** Not special type.

ACKNOWLEDGMENTS

Authors thank all participants for their contribution and cooperation throughout the study.

AUTHORS' CONTRIBUTIONS

HG: conception and design of the study. IA: design of the data collection tool, data collection, and revising the manuscript MM: design the data collection tool, data collection, analysis and interpretation, and writing the original draft. MA: design of the study and analysis and interpretation of data. All authors have read and approved the manuscript.

CONFLICT OF INTERESTS

There are no conflicts of interest.

REFERENCES

1. Ghoncheh M, Pournamdar Z, Salehiniya H. Incidence and mortality and epidemiology of breast cancer in the world. Asian Pacific Journal of Cancer Prevention [Internet]. 2016 Jun 1;17(sup3):43–6. Available from: https://doi.org/10.7314/apjcp.2016.17.s3.43

2. Dhadiala SK, Patankar S. Assessment of breast carcinoma by correlating BI-RADS scoring with mammographic density. International Surgery Journal [Internet]. 2020 Oct 23;7(11):3674. Available from: https://doi.org/10.18203/2349-2902.isj20204671

3. Sickles EA. ACR Appropriateness Criteria Breast Cancer Screening. Breast Dis Year Book Q. 2013;24(3):233–4. doi:10.1016/j.breastdis.2013.07.011

4. Brem RF, Tabár L, Duffy SW, Inciardi MF, Guingrich JA, Hashimoto BE, *et al*. Assessing Improvement in Detection of Breast Cancer with Three-dimensional Automated Breast US in Women with Dense Breast Tissue: The SomoInsight Study. Radiology [Internet]. 2014 Oct 20;274(3):663–73. Available from: https://doi.org/10.1148/radiol.14132832

5. Kutluer N, Aksu A, Bozan M, Kanat B, Kargici H, Cay F, *et al*. Correlation between histopathological results and BI-RADS classification in breast masses. Annals of Medical Research [Internet]. 2019 Jan 1;26(11):2698. Available from: https://doi.org/10.5455/annalsmedres.2019.07.366

6. BN N, Thomas S, Hiremath R, Alva SR. Comparison of diagnostic accuracy of BIRADS score with pathologic findings in breast lumps. Ann Pathol Lab Med. 2017;4(3):A236–42. Available from: https://www.pacificjournals.com/journal/index.php/apalm/article

7. Aziz S, Mohamad MA, Zin RRM. Histopathological correlation of breast carcinoma with Breast Imaging-Reporting and Data System. Malays J Med Sci [Internet]. 2022 Aug 26;29(4):65–74. Available from: https://doi.org/10.21315/mjms2022.29.4.7

8. Autier P, Boniol M. Mammography screening: a major issue in medicine. Eur J Cancer. 2018;90:34–62.