ORIGINAL ARTICLE
https://doi.org/10.29289/2594539420250038

Quaternary prevention in women's health: the importance of stratification of BI-RADS 4 breast lesions

Camila Caminha Kondlatsch[1]*, Guilherme Felipe Staudt[2], Regina Arruda Leal[2], Daniella Serafin Couto Vieira[3]

ABSTRACT

Introduction: Breast cancer is the most common malignant neoplasm among women in Brazil and worldwide, excluding non-melanoma skin tumors. The definitive diagnosis is based on the histopathological analysis of samples obtained through biopsies, performed after the identification of suspicious radiological findings on breast imaging studies. In this context, the Breast Imaging Reporting & Data System (BI-RADS) plays a central role in stratifying the risk of malignancy, particularly in the BI-RADS 4 category, which encompasses lesions with a broad spectrum of diagnostic probability and significant heterogeneity regarding histopathological outcomes. **Objective:** To evaluate the diagnostic performance of BI-RADS category 4 stratification through the correlation between radiological findings and histopathological results of breast biopsies. **Methods:** This is a descriptive, cross-sectional study based on the analysis of secondary data from a tertiary hospital in Southern Brazil for the year 2023. The variables analyzed were age, sex, BI-RADS category, sample type, lesion side, radiological finding, disease type, presence of histological microcalcifications, Nottingham histological grade, expression of tumor biomarkers, cell proliferation index, and immunophenotypic subtype. **Results:** Of the imaging findings analyzed, 14.9% were classified as BI-RADS 4 without subdivision. The diagnostic test performed without stratification had an accuracy of 29.87% (95% confidence interval [CI] 18.18%–43.86%), while the accuracy of the group with stratification of imaging findings was significantly higher, reaching 47.68% (95%CI 38.71%–56.76%). **Conclusions:** Stratification of the BI-RADS 4 category allowed for a more precise correlation between radiological suspicion and histopathological findings, resulting in greater diagnostic accuracy in biopsy indications. This process aligns with the concept of quaternary prevention, contributing to more assertive clinical management and a reduction in unnecessary interventions.

KEYWORDS: mammography; early cancer detection; breast neoplasm; quaternary prevention.

INTRODUCTION

Breast cancer is the most common malignant neoplasm among women in Brazil and worldwide[1]. By 2050, it is estimated that the number of new cases and deaths will increase by 38% and 68%, respectively[2]. The diagnosis of this disease is based on histopathological examination of material obtained from surgical biopsy or core needle biopsy, usually prompted by a radiological finding suggestive of malignancy[3,4].

Among the main breast imaging tests, mammography is widely used for population-based screening, aiming at the early diagnosis of breast cancer[5]. Until 2025, Brazil's Ministry of Health recommended breast cancer screening for women aged 50 to 69 with average risk[6]. Starting in September of that year, the guideline was expanded to include women up to age 74, with screening also permitted for those aged 40 to 49 based on individual clinical assessment[7]. Mammograms must be reported according

Institution where the study was conducted: Hospital Universitário Polydoro Ernani de São Thiago, Florianópolis (SC), Brazil.
[1]Universidade Federal de Santa Catarina, School of Medicine – Florianópolis (SC), Brazil.
[2]Hospital Universitário Polydoro Ernani de São Thiago, Hemotherapy, Hematology, and Oncology Unit – Florianópolis (SC), Brazil.
[3]Universidade Federal de Santa Catarina, Center for Health Sciences, Department of Pathology – Florianópolis (SC), Brazil.
***Corresponding author:** camila.c.k@hotmail.com
Editor: René Aloisio da Costa Vieira
Received on: 09/01/2025. **Accepted on:** 03/25/2026. **Published on:** 06/22/2026.
How to cite: Kondlatsch CC, Staudt GF, Leal RA, Vieira DSC. Quaternary prevention in women's health: the importance of stratification of BI-RADS 4 breast lesions. Mastology. 2026;36:e20250038. https://doi.org/10.29289/2594539420250038

to the Breast Imaging-Reporting and Data System (BI-RADS), a classification system proposed by the American College of Radiology (ACR) to standardize the results of breast imaging exams, thereby reducing ambiguity in clinical management and facilitating communication among professionals. In addition to mammography, breast ultrasound and magnetic resonance imaging (MRI) exams can also be classified into BI-RADS categories[8].

In the literature, mammography has moderate sensitivity, with reported values of approximately 54.5% (range: 27.0%–86.8%) and specificity of around 85.5% (range: 62.9%–98.8%), with lower performance in younger women and those with dense breasts. Ultrasound, in turn, has variable sensitivity, estimated at 67.2% (range: 26.9%–87.5%), and specificity of approximately 76.8% (range: 18.8%–96.9%), and is recommended as a complementary method. MRI stands out for its high sensitivity, estimated at 94.6% (range: 85.7%–100%), although it has more limited specificity, around 74.2% (range: 25%–100%)[9].

The BI-RADS Atlas classifies imaging findings into categories 0 through 6. Exams classified as BI-RADS 4 and 5 should be followed by biopsy, preferably fine-needle aspiration[3,4], of the lesion(s) and histopathological examination of the collected samples. Çoraplı et al. found positive predictive values (PPV) for lesions classified as BI-RADS 3, 4, and 5 on ultrasound of 2.15%, 47.44%, and 95.19%, respectively. On mammography, the corresponding PPVs were 3.79%, 53.45%, and 94.20%. Finally, on MRI, they were 0%, 57.89%, and 88.10%, respectively[10]. In mammography examinations, the stratification of the BI-RADS 4 category into subcategories showed a PPV of 7.6% for subcategory 4A, 22.0% for 4B, and 69.3% for 4C, as described by Elezaby et al[11]. In ultrasound, PPVs also increase within category 4[12]. Finally, although not formally recommended in MRI, the subdivision of the BI-RADS 4 category has proven useful by demonstrating increasing PPV across the categories[13].

Although the BI-RADS Atlas provides standardized descriptors and risk ranges, it does not define formal rules linking individual imaging features to specific categories. This limitation may contribute to an interpretation that is more examiner-dependent and involves a variable degree of subjectivity. In this context, due to the priority given to diagnostic sensitivity in population-based screening settings, the classification of findings as BI-RADS 4 tends to incorporate criteria with lower specificity[14], with an estimated 70%–80% of findings classified as such corresponding to benign lesions[15]. This high rate of false-positive results can lead to iatrogenic interventions, including the indication of potentially unnecessary biopsies[14,15].

Iatrogenesis is defined as harm caused to the patient as a result of medical interventions[16]. In this sense, quaternary prevention refers to actions aimed at preventing iatrogenic practices by protecting patients from medical interventions whose risks outweigh the benefits[17]. The lack of correlation between imaging and the histopathology of breast lesions can lead to iatrogenic practices, placing a burden on the healthcare system, subjecting women to invasive procedures, and causing psychological distress[18]. In view of this, the present study aimed to analyze the correlation between the BI-RADS category and the histopathological diagnoses of breast lesions, with an emphasis on the stratification of BI-RADS category 4, in line with the principles of quaternary prevention and the minimization of iatrogenic complications.

METHODS

This is an observational, descriptive, cross-sectional, and retrospective study based on the analysis of secondary data from a tertiary hospital in Southern Brazil in 2023. The objective was to compare the BI-RADS category and the risk of malignancy attributed to imaging findings with the corresponding histopathological reports, to assess whether there was a radiological and pathological correlation among the breast lesions studied. The analysis was conducted by comparing a non-stratified model, in which the BI-RADS 4 category was evaluated globally, and a stratified model, based on the subdivision into BI-RADS 4A, 4B, and 4C. The study included patients from the Brazilian Ministry of Health mammography screening program (women aged 50 to 69 years) who underwent core needle biopsy in 2023. Information regarding the clinical indication for imaging exams (screening *versus* diagnostic evaluation) was not available in the institutional database used in this study, as the dataset included imaging reports but did not systematically record the original requests for the exams. The following were excluded: males, individuals under 50 years or 70 years of age or older, fine-needle aspiration samples, breast surgical specimens, axillary lymph nodes from mastectomies with sentinel lymph node biopsy and/or axillary dissection, and excisional biopsies. The selection process is shown in Figure 1.

Data collection was conducted entirely in a hospital setting and *online*. Initially, the Information System of the National Breast

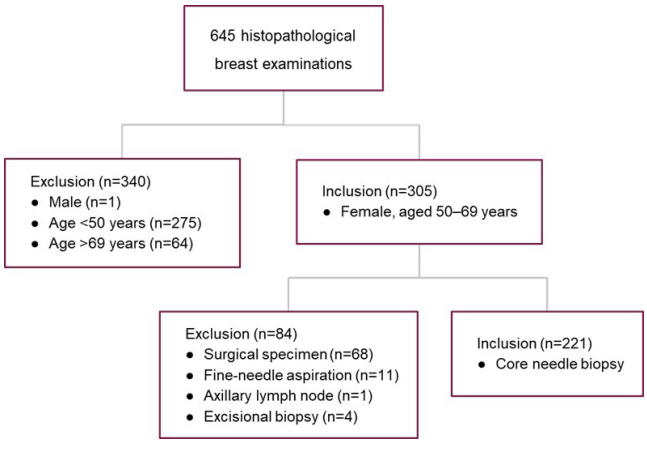

Figure 1. Inclusion and exclusion criteria.

Cancer Control Program (SISMAMA) was used, from which a list of patients who underwent breast histopathological analysis at this facility in 2023 was obtained. These patients were referred to the facility through the National Referral System. Breast imaging examinations prompting biopsy included mammography, ultrasound, and MRI. All imaging exams were interpreted by radiologists certified by the Brazilian College of Radiology and Diagnostic Imaging (CBR), regardless of the originating institution. Multiple radiologists participated in the process, which made it impossible to standardize the reports and assess interobserver agreement. Finally, the Pathological Anatomy Laboratory Unit System (SULAP) was used to obtain the histopathological reports analyzed in the study.

The variables collected were:

- BI-RADS category provided by the imaging exam (mammography, ultrasound, or MRI): 0, 1, 2, 3, 4, 4A, 4B, 4C, 5, or 6;
- Lateralization of the lesion: right or left;
- Radiological findings: mass, microcalcifications, distortion, or asymmetry;
- Presence of histological microcalcifications;
- Type of disease: benign or malignant;
- Histological subtype: invasive ductal carcinoma, unspecified (IDC-US), ductal carcinoma in situ (DCIS), or invasive lobular carcinoma (ILC);
- Nottingham histological grade: G1, G2, or G3;
- Presence of estrogen receptors (ER);
- Presence of progesterone receptors (PR);
- Presence of human epidermal growth factor receptor 2 (HER2): immunohistochemistry (IHC) 0, IHC 1+, IHC 2+, IHC 3+;
- Expression of the Ki-67 marker in immunohistochemical analysis: <15% or ≥15%;
- Immunophenotypic subtype: Luminal A, Luminal B HER2-, Luminal B HER2+, HER2-positive, or triple-negative (TNBC).

For the descriptive analysis of the data, the absolute and relative frequencies of the variables were estimated and were compared between groups and with the current literature. The frequency assessments of the variables were described in tables, rounded to the first decimal place. Sensitivity, specificity, predictive values, and accuracy were calculated using MedCalc software. The discriminatory power of the methods was assessed using the receiver operating characteristic (ROC) curve and the area under the curve (AUC), estimated in Jamovi software. The estimates were presented with their respective 95% confidence intervals (95%CI).

The study was approved by the Research Ethics Committee under registration number CAAE 78604224.1.0000.0121 and opinion number 6.951.785, in accordance with Resolution No. 466 of December 12, 2012[19].

RESULTS

Distribution of diagnoses

MRI scans were excluded from the analysis in accordance with the exclusion criteria previously defined in the study protocol. Thus, the final sample included only mammograms and ultrasound scans, comprising 186 (84.2%) mammograms and 35 (15.8%) ultrasound scans.

Of the total cases, 151 (68.3%) were benign, and 59 (26.7%) were diagnosed with breast cancer; 11 (4.9%) histopathological examinations had an inconclusive diagnosis.

As shown in Table 1, Mass is the most prevalent imaging finding in both groups, followed by microcalcifications, which were more common in benign cases (45.0%) than in malignant cases (20.3%). Table 2 presents the distribution of lesions according to BI-RADS category.

Benign lesions were concentrated in reports with lower BI-RADS categories: 3, 4, 4A, and 4B. BI-RADS 4C and 5 were more prevalent in diagnoses of malignancy. The majority (76.8%) of reports with BI-RADS category 4 were stratified. Finally, in 22.5% of benign cases and 12.6% of malignant cases, it was not possible to determine the BI-RADS category. A total of 77 (41.4%) mammograms showed BI-RADS categories highly suggestive of breast cancer (categories 4B, 4C, and 5). For ultrasound, this rate was lower, with 10 (28.6%) cases highly suggestive of the disease.

Malignant histopathological diagnoses

Breast tumors diagnosed as DCIS accounted for 11 cases (16.9%); 80% of the imaging findings leading to this diagnosis were microcalcifications.

Table 1. Characterization of benign and malignant diagnoses.

	Benign n (%)	Malignant n (%)
Lateralization		
Right	67 (44.4)	32 (54.2)
Left	84 (55.6)	27 (45.7)
Image finding		
Mass	73 (48.3)	42 (71.2)
Distortion	7 (4.6)	5 (8.5)
Microcalcifications	68 (45.0)	12 (20.3)
Asymmetry	3 (1.9)	0 (0.0)
Histological microcalcifications		
Present	46 (30.5)	14 (23.7)
Absent	32 (21.2)	40 (67.8)
Not listed	73 (48.3)	5 (8.5)
Total	151 (100)	59 (100)

Malignancy diagnoses excluding DCIS comprised 45 (91.8%) cases of the DCIS not otherwise specified (NOS) type. In contrast, ILC affected 4 (8.16%) patients in this group.

The immunohistochemical profile of malignant tumors, excluding DCIS, is presented in Table 3.

High positivity was observed for hormone receptors, especially for ER, while the HER2 marker was predominantly negative. Ki-67 expression showed a similar distribution between low and high levels. The majority (75%) of tumors with high Ki-67 expression were classified as BI-RADS 4C and 5.

As demonstrated in Table 4, among malignant cases, a predominance of luminal-subtype tumors was observed, particularly Luminal A (40.8%) and Luminal B (32.6%). The other molecular subtypes occurred less frequently, including TNBC (10.2%) and pure HER2 (2.0%) tumors. Regarding histological grade, there was a higher prevalence of Nottingham G2 tumors (46.9%), followed by G3 (28.6%) and G1 (24.5%). It was also observed that 6 (40.0%) high-grade tumors were classified as BI-RADS 4C and 5.

Radiological and pathological correlation

Tables 5 and 6 evaluated the radiological performance of the imaging studies analyzed. The prevalence of breast cancer in the female population was estimated at 12.5%[20].

In the diagnostic test shown in Table 5, BI-RADS 4B, 4C, and 5 were considered positive tests, and BI-RADS 3 and 4A were considered negative tests. Out of a total of 126 examinations, 29 (23.0%) true positives and 39 (30.9%) true negatives were obtained, in addition to 4 (3.2%) false negatives and 54 (42.8%) false positives. The test had a sensitivity of 87.88% (95%CI 71.80%–96.60%) and a specificity of 41.94% (95%CI 31.78%–52.62%). The positive likelihood ratio was 1.51 (95%CI 1.22–1.88) and the negative likelihood ratio was 0.29 (95%CI 0.11–0.75)[21]. The PPV was calculated as PPV3, defined as the proportion of malignant diagnoses among biopsies actually performed[22], and resulted in 17.78% (95%CI 14.86%–21.13%). The negative predictive value (NPV) was 96.03 (95%CI 90.36%–98.43%). The test had an accuracy of 47.68% (95%CI 38.71%–56.76%)[20]. ROC curve analysis demonstrated moderate discriminatory ability, with an AUC of 0.76 (95%CI 0.62–0.89; p<0.001)[23].

When BI-RADS 4 imaging reports without stratification and BI-RADS 5 reports were considered positive tests, and BI-RADS

Table 3. Characterization of tumor markers in malignant diagnoses, excluding ductal carcinoma *in situ*.

Tumor marker	Frequency n (%)
RE	
Positive	43 (87.7)
Negative	6 (12.2)
RP	
Positive	29 (59.2)
Negative	20 (40.8)
HER2	
IHC 0	17 (34.7)
IHQ 1+	23 (46.9)
IHQ 2+	5 (10.2)
IHQ 3+	4 (8.2)
Ki-67	
High (≥15%)	25 (51.0)
Low (<15%)	24 (49.0)
Total	49 (100)

ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal growth factor receptor 2; IHC: immunohistochemistry.

Table 4. Distribution of Nottingham histological grades of malignant diagnoses other than ductal carcinoma *in situ* by immunophenotypic subtype.

	G1 n (%)	G2 n (%)	G3 n (%)	Total n (%)
Luminal A	12 (24.5)	8 (16.3)	0 (0.0)	20 (40.8)
Luminal B	0 (0.0)	8 (16.3)	8 (16.3)	16 (32.6)
Luminal B HER2+	0 (0.0)	2 (4.0)	1 (2.0)	3 (6.1)
HER2-positive	0 (0.0)	1 (2.0)	0 (0.0)	1 (2.0)
TNBC	0 (0.0)	1 (2.0)	4 (8.2)	5 (10.2)
Undetermined	0 (0.0)	3 (6.1)	1 (2.0)	4 (8.2)
Total	12 (24.5)	23 (46.9)	14 (28.6)	49 (100)

HER2: human epidermal growth factor receptor 2; TNBC: triple-negative breast cancer.

Table 2. Distribution of benign and malignant histopathological diagnoses by BI-RADS.

Diagnosis	BI-RADS							
	3 n (%)	4 n (%)	4A n (%)	4B n (%)	4C n (%)	5 n (%)	Not listed n (%)	Total n (%)
Benign	7 (3.4)	24 (11.4)	32 (15.2)	42 (20.0)	10 (4.7)	2 (0.9)	34 (16.1)	151 (71.9)
Malignant	2 (0.9)	7 (3.4)	2 (0.9)	6 (2.8)	11 (5.2)	12 (5.7)	19 (9.0)	59 (28.0)
Total	9 (4.2)	31 (14.7)	34 (16.1)	48 (22.8)	21 (10.0)	14 (6.7)	53 (25.2)	210 (100)

3 reports were considered negative tests (as per Table 5), the sensitivity was 90.48% (95%CI 69.62%–98.83%) and the specificity was 21.21% (95%CI 8.98%–38.91%). The positive likelihood ratio was 1.15 (95%CI 0.92%–1.44) and the negative likelihood ratio was 0.45 (95%CI 0.10–1.96). A PPV3 of 14.09% (95%CI 11.58%–17.04%) and an NPV of 93.97% (95%CI 78.13%–98.55%) were obtained. The test achieved 29.87% (95%CI 18.18%–43.86%) accuracy[20]. The ROC curve showed an AUC of 0.80 (95%CI 0.70–0.90; p<0.001)[23].

DISCUSSION

The study population was defined as the target group for breast cancer screening in Brazil according to the current public health guidelines in effect at the time of data collection (women aged 50 to 69 years)[6]. It was observed that more than half (52.7%) of the histopathological investigations were performed on individuals outside the age range recommended by the Ministry of Health for breast cancer screening at the time. This finding suggests a possible expansion of diagnostic indications beyond the recommended screening, in line with evidence that women under 50 years of age should be included in mammographic screening, given the prevalence of breast cancer in this population in Brazil[24]. Furthermore, the data highlight the study's limitation in systematically evaluating the clinical and care context of these investigations outside the screening setting.

According to the Ministry of Health's new guideline, screening mammography is recommended up to age 74, with subsequent management being individualized[7]. Expanding population-based screening to older age groups may increase the risk of overdiagnosis, characterized by the identification of neoplasms that would not progress to clinical consequences over the course of a woman's life[25]. This phenomenon often results from screening examinations or the broadening of diagnostic criteria and can lead to unnecessary treatments, side effects, anxiety, stigmatization, and reduced effectiveness of the healthcare system[25,26]. In the context of breast cancer, overdiagnosis occurs in 31.0% of women aged 70 to 74 who undergo screening mammography and in 47.0% of women aged 75 to 84 who participate in *screening*[27].

Regarding the correlation between BI-RADS categories and histopathological diagnoses, BI-RADS 4 is frequently assigned in numerous cases. The use of category 4 without stratification has been considered inappropriate since the 4th edition of the BI-RADS Atlas[8]. The continued failure to adhere to the subdivision of the BI-RADS 4 category points to an opportunity to improve the quality of mammographic interpretation. Another important finding is the biopsies performed for BI-RADS 3 lesions, a practice not supported by the guidelines[8]. Finally, the BI-RADS category could not be determined in 24.4% of the imaging exams analyzed, leading to the exclusion of these cases from the diagnostic performance evaluation. This limitation reflects deficiencies in the completion of medical records, resulting in the loss of information relevant to oncological investigation.

The high proportion of mammograms classified as highly suspicious for malignancy (BI-RADS 4B, 4C, and 5), observed in 41.3% of cases, contrasts with what is described in population-based screening settings, where these categories generally account for about 2.5% of the exams performed[11]. This finding can be explained by the care profile of the service studied, a tertiary hospital that tends to concentrate patients previously selected based on suspicious clinical or radiological findings, thereby increasing the pre-test probability of malignancy. On ultrasound, the proportion of highly suspicious lesions was lower (28.5%), which may reflect differences inherent to the indications and the complementary role of this method in the evaluation of breast lesions[28].

Ki-67 marker expression showed a similar distribution between high- and low-expression tumors, with the majority (75.0%) of high-expression tumors classified as BI-RADS 4C and 5. The association between high Ki-67 scores and higher BI-RADS categories has previously been described in the literature, demonstrating a strong positive correlation between these parameters[29].

The prevalence of immunophenotypic subtypes found in the present study was similar to that reported in the literature, considering the age range of the screened population[30]. Cases classified as inconclusive resulted from indeterminate HER2 status (IHC 2+), due to the unavailability of the fluorescent *in situ* hybridization (FISH) test for confirmatory analysis, as recommended by the College of American Pathologists[31].

A correlation was identified between Nottingham low-grade (G1) tumors and the more indolent immunophenotype (Luminal A). In contrast, TNBCs, the more aggressive subtype, predominated as

Table 5. Positive predictive value of highly suspicious imaging findings for breast cancer (BI-RADS 4B, 4C, and 5).

		Disease – Malignancy n (%)		Total
		Present	Absent	
Test BI-RADS	Positive (4B, 4C, and 5)	29 (23.0)	54 (42.8)	83 (65.9)
	Negative (3 and 4A)	4 (3.2)	39 (30.9)	43 (34.1)
	Total	33 (26.2)	93 (73.8)	126 (100)

Table 6. Positive predictive value of imaging findings classified as BI-RADS 4 without stratification.

		Disease – Malignancy n (%)		Total
		Present	Absent	
Test BI-RADS	Positive (4 and 5)	19 (35.2)	26 (48.1)	45 (83.4)
	Negative (3)	2 (3.7)	7 (12.9)	9 (16.7)
	Total	21 (38.9)	33 (61.2)	54 (100)

poorly differentiated tumors (G3). Among high-grade tumors, 40.0% were classified as BI-RADS 4C and 5. The integration of imaging findings with tumor biology represents a relevant field for future research, with the potential to improve lesion characterization and advance more personalized diagnostic and therapeutic approaches.

Regarding the analysis of the radiological-pathological correlation in the stratified test, there was a high rate of false-positive results, which implies the potential performance of iatrogenic breast biopsies. The false-positive result is the primary harm associated with screening mammography. It is estimated that the risk of such a result in screening exams is approximately 17.0%, and about 2.9% of women undergo biopsy thereafter[32]. The stratified test achieved a sensitivity of 87.88%, considered satisfactory for a screening test targeting a sensitivity of over 75.0%[22]. Specificity, however, was below the acceptable range: 41.94%, with 88%–95% considered satisfactory[22]. Moreover, the accuracy of 47.68% falls well short of the ideal. The accuracy of a screening test is crucial, as erroneous results can have profound repercussions on patients' health. False-positive results impact psychological well-being as well as physical integrity, since biopsy is an invasive procedure with inherent risks[18]. Furthermore, considering the limited resources available in the Brazilian Unified Health System[33], it can be inferred that an inappropriately indicated biopsy may delay or prevent the procedure from being performed on a patient who truly needs the test. On the other hand, false-negative results may delay diagnosis and consequently lead to a worse prognosis[31].

The analysis in Table 6 sought to identify whether there was a change in the radiological and pathological correlation when BI-RADS 4 stratification was not used. Without stratification, there was a slight increase in sensitivity, but a drop of more than 20 percentage points in specificity. The PPV3 and NPV decreased compared to the stratification group. The PPV3 observed in this study, whether in the stratified group (17.78%) or the non-stratified group (14.09%), was lower than that reported by the Breast Cancer Surveillance Consortium (28.60%)[21]. This finding may reflect the adoption of a radiological interpretation approach with greater emphasis on sensitivity, resulting in biopsies being performed on lesions with a lower pre-test probability of malignancy. Although this strategy reduces the risk of false-positive results, it may simultaneously contravene the principles of quaternary prevention by increasing the proportion of unnecessary interventions. In addition, performing biopsies on BI-RADS 3 lesions, considered inappropriate, may have contributed to the reduction in the PPV3. Accuracy also dropped significantly: from 47.68% in the stratified group to 29.87% in the unstratified group. An even higher AUC was observed in the group subjected to BI-RADS stratification compared to the unstratified group, suggesting better discriminatory power, although no formal statistical comparison was performed between the curves.

Some limitations inherent to the study design should be considered, such as its retrospective nature and the use of secondary data. The unavailability of approximately one-quarter of the BI-RADS categories in the sample may have partially influenced the interpretation of the results. Furthermore, the evaluations were performed by a single observer, without an interobserver agreement analysis, which may limit the reproducibility of the findings. The composition of the sample should also be interpreted with caution, as it includes patients treated at a tertiary hospital, a setting that tends to concentrate cases with higher clinical and radiological suspicion. Additionally, it was not possible to distinguish between exams performed in the context of screening and those for diagnostic purposes, as this information was not available in the institutional database. These factors may have influenced the distribution of BI-RADS categories and the observed prevalence of malignancy, limiting the generalizability of the results to other populations and screening models. Despite these limitations, the study offers a relevant analysis of breast lesions classified as BI-RADS 4 in a real-world clinical setting, contributing to the understanding of their anatomopathological correlation.

CONCLUSION

In the present study, the evaluation of diagnostic performance without stratification of the BI-RADS 4 category demonstrated limited accuracy in the correlation between radiological findings and histopathological results. On the other hand, the subclassification of this category was associated with greater agreement between the radiological subgroups and the anatomopathological outcomes. These findings indicate that the stratification of BI-RADS 4 lesions allows for a more accurate estimation of the risk of malignancy, favoring greater alignment between radiological suspicion and histological diagnosis. In this context, the evaluated strategy demonstrates potential to improve the risk stratification of breast lesions and more accurately guide the indication for biopsies in clinical practice, although additional studies are needed for its validation in different settings.

ACKNOWLEDGEMENT

We would like to thank the staff of the Department of Pathological Anatomy and the team involved in oncology care and breast imaging at Polydoro Ernani de São Thiago University Hospital for their institutional support and assistance in providing access to the data used in this study.

AUTHORS' CONTRIBUTIONS

CCK: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. GFS: Formal analysis, Validation, Writing – review & editing. RAL: Formal analysis, Validation, Writing – review & editing. DSCV: Conceptualization, Methodology, Project administration, Supervision, Validation, Writing – review & editing.

Funding: none.

Conflict of interests: nothing to declare.

Artificial intelligence usage: ChatGPT - Language editing and improvement of clarity and fluency. I declare that any use of generative AI was appropriately supervised by the authors, transparently disclosed in the manuscript, and does not confer authorship to the tool. I declare that the authors retain full responsibility for the content of the manuscript.

Data availability statement: The data underlying this research were obtained from medical records and are protected by the ethical and legal confidentiality standards in force in Brazil (Law No. 13,709/2018, CNS Resolutions No. 466/2012 and No. 510/2016, and CFM Resolution No. 2,217/2018). However, the data may be provided upon formal request to the corresponding author, with a clear description of the intended use, a commitment to maintaining confidentiality, a pledge not to attempt to identify participants, and the signing of a confidentiality agreement, in accordance with institutional policies.

REFERENCES

1. Miranda Filho A, Laversanne M, Ferlay J, Colombet M, Piñeros M, Znaor A, et al. The GLOBOCAN 2022 cancer estimates: Data sources, methods, and a snapshot of the cancer burden worldwide. Int J Cancer. 2025;156(7):1336-46. https://doi.org/10.1002/ijc.35278
2. Kim J, Harper A, McCormack V, Houssami N, Morgan E, Mutebi M, et al. Global patterns and trends in breast cancer incidence and mortality across 185 countries. Nat Med. 2025;31(4):1154-62. https://doi.org/10.1038/s41591-025-03502-3
3. Bruening W, Fontanarosa J, Tipton K, Treadwell J, Launders J, Schoelles K. Systematic Review: Comparative Effectiveness of Core-Needle and Open Surgical Biopsy to Diagnose Breast Lesions. Ann Inter Med. 2010;152(4):238-46. https://doi.org/10.7326/0003-4819-152-1-201001050-00190
4. Shanmugalingam A, Hitos K, Hegde S, Al-Mashat A, Pathmanathan N, Edirimmane S, et al. Concordance between core needle biopsy and surgical excision for breast cancer tumor grade and biomarkers. Breast Cancer Res Treat. 2022;193(1):151-9. https://doi.org/10.1007/s10549-022-06548-w
5. Shi J, Li J, Gao Y, Chen W, Zhao L, Li N, et al. The screening value of mammography for breast cancer: an overview of 28 systematic reviews with evidence mapping. J Cancer Res Clin Oncol. 2025;151(3):102. https://doi.org/10.1007/s00432-025-06122-z
6. Ren W, Chen M, Qiao Y, Zhao F. Global guidelines for breast cancer screening: A systematic review. Breast. 2022;64:85-99. https://doi.org/10.1016/j.breast.2022.04.003
7. Brasil. Ministério da Saúde. Secretaria de Atenção Especializada à Saúde. Departamento de Atenção Especializada e Temática. Coordenação-Geral de Atenção às Doenças Crônicas. Nota Técnica nº 626/2025: orientações sobre o rastreamento do câncer de mama no âmbito do Sistema Único de Saúde [Internet]. Brasília (DF): Ministério da Saúde; 2025 [cited on 2026 Mar 21]. Available from: https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/notas-tecnicas/2025/nota-tecnica-no-626-2025-cgcan-decan-saes-ms.pdf
8. Newell MS, Destounis SV, Leung JWT, DeMartini WB, Lee CH, Eby PR. ACR Breast Imaging Reporting and Data System: Manual. Reston: American College of Radiology; 2025.
9. Aristokli N, Polycarpou I, Themistocleous SC, Sophocleous D, Mamais I. Comparison of the diagnostic performance of Magnetic Resonance Imaging (MRI), ultrasound and mammography for detection of breast cancer based on tumor type, breast density and patient's history: A review. Radiography (Lond). 2022;28(3):848-56. https://doi.org/10.1016/j.radi.2022.01.006
10. Çoraplı M, Bulut HT, Çoraplı G, Pehlivanoğlu B, Alakuş H, Soysaldı H. Likelihood of cancer in breast cancer imaging according to BI-RADS. J Surg Med. 2021;5(9):857-61. https://doi.org/10.28982/josam.958518
11. Elezaby M, Li G, Bhargavan-Chatfield M, Burnside ES, DeMartini WB. ACR BI-RADS Assessment Category 4 Subdivisions in Diagnostic Mammography: Utilization and Outcomes in the National Mammography Database. Radiology. 2018;287(2):416-22. https://doi.org/10.1148/radiol.2017170770
12. He P, Cui LG, Chen W, Yang RL. Subcategorization of Ultrasonographic BI-RADS Category 4: Assessment of Diagnostic Accuracy in Diagnosing Breast Lesions and Influence of Clinical Factors on Positive Predictive Value. Ultrasound Med Biol. 2019;45(5):1253-8. https://doi.org/10.1016/j.ultrasmedbio.2018.12.008
13. Strigel RM, Burnside ES, Elezaby M, Fowler AM, Kelcz F, Salkowski LR, et al. Utility of BI-RADS Assessment Category 4 Subdivisions for Screening Breast MRI. AJR Am J Roentgenol. 2017;208(6):1392-9. https://doi.org/10.2214/ajr.16.16730
14. Vatteroni G, Dietzel M, Baltzer PAT. Can structured integration of BI-RADS criteria by a clinical decision rule reduce the number of unnecessary biopsies in BI-RADS 4 lesions? A systematic review and meta-analysis. Eur Radiol. 2025;35(3):1504-13. https://doi.org/10.1007/s00330-024-11274-6
15. Liu C, Sun M, Arefan D, Zuley M, Sumkin J, Wu S. Deep learning of mammogram images to reduce unnecessary breast biopsies: a preliminary study. Breast Cancer Res. 2024;26(1):82. https://doi.org/10.1186/s13058-024-01830-9
16. Yang J, Zhou S, Feng X, Chen Y, Hu Y, Xu M. Understanding the Effects of Iatrogenic Management on Population Health: A Medical Innovation Perspective. China CDC Wkly. 2023;5(27):614-8. https://doi.org/10.46234/ccdcw2023.118
17. Martins C, Godycki-Cwirko M, Heleno B, Brodersen J. Quaternary prevention: reviewing the concept. Eur J Gen Pract. 2018;24(1):106-11. https://doi.org/10.1080/13814788.2017.1422177
18. Canelo-Aybar C, Ferreira DS, Ballesteros M, Posso M, Montero N, Solà I, et al. Benefits and harms of breast cancer mammography screening for women at average risk of breast cancer: A systematic review for the European Commission Initiative on Breast Cancer. J Med Screen. 2021;28(4):389-404. https://doi.org/10.1177/0969141321993866

19. Brasil. Conselho Nacional de Saúde. Resolução nº 466, de 12 de dezembro de 2012. Diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos [Internet]. Brasília: Diário Oficial da União; 2012 [cited on 2025 Feb 13]. Available from: https://bvsms.saude.gov.br/bvs/saudelegis/cns/2013/res0466_12_12_2012.html.

20. Feuer E, Wun L, Boring C, Flanders W, Timmel M, Tong T. The lifetime risk of developing breast cancer. J Natl Cancer Inst. 1993;85(11):892-7. https://doi.org/10.1093/JNCI/85.11.892

21. MedCalc Software Ltd. MedCalc Statistical Software [Internet]. Version 23.0.8. Belgium: MedCalc; 2024 [cited on 2025 Mar 9]. Available from: https://www.medcalc.org.

22. Lehman CD, Arao RF, Sprague BL, Lee JM, Buist DS, Kerlikowske K, et al. National Performance Benchmarks for Modern Screening Digital Mammography: Update from the Breast Cancer Surveillance Consortium. Radiology. 2017;283(1):49-58. https://doi.org/10.1148/radiol.2016161174

23. The Jamovi Project. ***jamovi*** [Computer Software]. Version 2.7. Sydney: The jamovi project; 2025 [cited on 2026 Jan 30]. Available from: https://www.jamovi.org.

24. Rosa DD, Bines J, Werutsky G, Barrios C, Cronemberger E, Queiroz GS, et al. The impact of sociodemographic factors and health insurance coverage in the diagnosis and clinicopathological characteristics of breast cancer in Brazil: AMAZONA III study (GBECAM 0115). Breast Cancer Res Treat. 2020;183(3):749-57. https://doi.org/10.1007/s10549-020-05831-y

25. Welch HG, Black WC. Overdiagnosis in cancer. J Natl Cancer Inst. 2010;102(9):605-13. https://doi.org/10.1093/jnci/djq099

26. Brodersen J, Schwartz LM, Heneghan C, O'Sullivan JW, Aronson JK, Woloshin S. Overdiagnosis: what it is and what it isn't. BMJ Evid Based Med. 2018;23(1):1-3. https://doi.org/10.1136/ebmed-2017-110886

27. Brawley OW, Ramalingam R. Understanding the Varying Biological Behaviors of Breast and Other Types of Cancer to Avoid Overdiagnosis. Ann Intern Med. 2023;176(9):1273-4. https://doi.org/10.7326/m23-1895

28. 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. https://doi.org/10.3390/cancers15072112

29. Ciurescu S, Cerbu S, Dima CN, Buciu V, Şerban DM, Ilaş DG, et al. Stratifying Breast Lesion Risk Using BI-RADS: A Correlative Study of Imaging and Histopathology. Medicina (Kaunas). 2025;61(7):1245. https://doi.org/10.3390/medicina61071245

30. Park S, Koo JS, Kim MS, Park HS, Lee JS, Lee JS, et al. Characteristics and outcomes according to molecular subtypes of breast cancer as classified by a panel of four biomarkers using immunohistochemistry. Breast. 2012;21(1):50-7. https://doi.org/10.1016/j.breast.2011.07.008

31. Wolff AC, Somerfield MR, Dowsett M, Hammond MEH, Hayes DF, McShane LM, et al. Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology–College of American Pathologists Guideline Update. Arch Pathol Lab Med. 2023;147(9):993-1000. https://doi.org/10.5858/arpa.2023-0950-SA

32. Løberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. Breast Cancer Res. 2015;17(1):63. https://doi.org/10.1186/s13058-015-0525-z

33. Castro MC, Massuda A, Almeida G, Menezes-Filho NA, Andrade MV, Souza Noronha KVM, et al. Brazil's unified health system: the first 30 years and prospects for the future. Lancet. 2019;394(10195):345-56. https://doi.org/10.1016/s0140-6736(19)31243-7