

Accuracy of mammographic findings in breast cancer: correlation between BI-RADS classification and histological findings*

Acurácia dos achados mamográficos do câncer de mama: correlação da classificação BI-RADS e achados histológicos

José Hermes Ribas do Nascimento¹, Vinícius Duval da Silva², Antônio Carlos Maciel³

Abstract **OBJECTIVE:** The present study was aimed at evaluating the BI-RADS® classification accuracy in mammography. Additionally, the frequency of different findings was described and the interobserver agreement was evaluated. **MATERIALS AND METHODS:** Mammographic images of 115 patients were independently and blindly reviewed by two specialists in compliance with BI-RADS recommendations, and later compared with histological data. The BI-RADS accuracy in mammography was evaluated. The interobserver agreement was analyzed with the Cohen's kappa (κ) test, and the differences between groups were evaluated with the chi-squared test. **RESULTS:** The present study demonstrated that the mammographic accuracy ranged from 75% to 62% in the differentiation between benign and malignant lesions with the utilization of the BI-RADS classification. Statistically significant interobserver agreement was observed in the description of masses margins ($\kappa = 0.66$). A low agreement rate was identified in the description of masses borders (shape) ($\kappa = 0.40$) and calcifications, both in relation to their distribution ($\kappa = 0.24$) and morphology ($\kappa = 0.36$). **CONCLUSION:** The present study demonstrated the BI-RADS accuracy in the differentiation between benign and malignant lesions. The interobserver agreement was poor in the analysis of calcifications morphology and distribution, but a progressive increase in the positive predictive values was observed in the subcategory 4. **Keywords:** Breast cancer; Mammography; Histopathology; Accuracy; BI-RADS; Ultrasonography.

Resumo **OBJETIVO:** A proposta deste estudo foi avaliar a acurácia da classificação BI-RADS® na mamografia. Os pontos secundários foram descrever a frequência de apresentação dos diferentes achados mamográficos e avaliar a concordância entre observadores. **MATERIAIS E MÉTODOS:** Os exames de 115 pacientes, encaminhados para *core biopsy*, foram reavaliados independentemente por dois médicos especialistas, cegados, utilizando a recomendação do BI-RADS. Posteriormente, os exames foram comparados com a histologia. A acurácia da classificação BI-RADS na mamografia foi avaliada. A concordância entre os médicos foi calculada pela estatística kappa (κ) de Cohen e as diferenças nos grupos de comparação foram analisadas com teste qui-quadrado. **RESULTADOS:** Esta pesquisa demonstrou que a acurácia mamográfica oscilou de 75% a 62% na diferenciação entre lesões benignas de malignas com o uso do BI-RADS. Houve importante concordância na descrição das margens dos nódulos ($\kappa = 0,66$). Baixa concordância foi identificada na descrição dos contornos (formas) dos nódulos ($\kappa = 0,40$) e na descrição das calcificações, tanto em relação à sua distribuição ($\kappa = 0,24$) como também em relação à morfologia ($\kappa = 0,36$). **CONCLUSÃO:** O presente estudo demonstrou que o método é acurado na diferenciação de lesões benignas de malignas. A concordância foi fraca na análise das calcificações quanto a morfologia e distribuição, no entanto, identificou-se elevação progressiva dos valores preditivos positivos nas subcategorias 4. **Unitermos:** Neoplasia da mama; Mamografia; Histologia; Acurácia; BI-RADS; Ultrassonografia.

Nascimento JHR, Silva VD, Maciel AC. Accuracy of mammographic findings in breast cancer: correlation between BI-RADS classification and histological findings. *Radiol Bras.* 2010;43(2):91–96.

INTRODUCTION

The Breast Imaging Reporting and Data System (BI-RADS®) developed by the American College of Radiology (ACR) was published in 1993 for mammography, and it was updated in 2003, with the intro-

duction of BI-RADS for ultrasonography and magnetic resonance imaging, with the objective of standardizing the assessment and reporting of breast lesions, and providing mastologists with guidance on the probability of malignancy of a given lesion by helping to conduct the investigation^(1,2), thus minimizing confusion in the images description and interpretation, and facilitating the elaboration of the final reports of

* Study developed at Clínica de Radiodiagnóstico Imagem Ltda., Santo Ângelo, RS, Brazil.

1. Master, MD, Radiologist, Director of the Clínica de Radiodiagnóstico Imagem Ltda., Professor at Division of Imaging Diagnosis, Instituto Cenecista de Ensino Superior Santo Ângelo, Santo Ângelo, RS, Brazil.

2. PhD, Associate Professor, Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, RS, Brazil.

3. PhD, MD, Radiologist at Hospital de Clínicas e Irmandade da Santa Casa de Misericórdia, Porto Alegre, RS, Brazil.

Mailing address: Dr. José Hermes Ribas do Nascimento. Rua Marechal Floriano, 774, Meller Sul. Santo Ângelo, RS, Brazil, 98801-650. E-mail: josehermesribas@hotmail.com

Received August 11, 2009. Accepted after revision January 11, 2010.

breast studies. This system comprises a specific vocabulary for describing each lesion and, as a report conclusion, the study result is classified into categories ranging from 0 to 6 according to the degree of suspicion of the findings, based on the positive predictive value (PPV) of the imaging study for breast cancer.

The BI-RADS is structured in four sections: section I – breast imaging lexicon; section II – reporting systematization; section III – follow-up and outcome monitoring; section IV – creation of a nationwide database⁽¹⁾.

A mammography is considered as negative for breast cancer when classified into BI-RADS categories 1, 2 and 3, and positive in the remaining categories. At category 1, there is no significant finding, the breasts are symmetrical, with no calcifications, masses, asymmetries, focal distortions or other alterations. At category 2, definitely benign findings are described, and at category 3 findings with < 2% chance of malignancy are described, with recommendation of a six-month follow-up evaluation. Category 0 corresponds to an incomplete study, requiring a complementary imaging study or even comparison with previous images. This is almost always recommended in a screening situation⁽²⁾.

Category 4 is reserved for those findings that do not present the classical malignancy appearance, but do present a wide spectrum of malignancy probability that is greater than that of category 3 lesions. According to BI-RADS, category 4 comprises lesions with a malignancy probability ranging from 3% to 94%, while those with a probability of 95% or more are classified as category 5. The approach recommended in category 4 cases is the request for cytological or histological investigation, while for cases in category 5, surgery is mandatory^(1,2).

The fourth BI-RADS edition was issued in 2003, and brought an update on the lesions descriptors (lexicon). The morphological description of microcalcifications was broken into the following categories that predict malignancy or benignity: a) typically benign; b) intermediate; c) high probability of malignancy⁽³⁾. Pleomorphic microcalcifications were subdivided into coarse heterogeneous (with intermediate

degree of concern) and fine pleomorphic linear (category with high malignancy probability)^(2,3). Heterogeneous microcalcifications are irregular, generally larger than 0.5 mm, and are considered as intermediate degree of concern, as well as amorphous or indistinct microcalcifications⁽³⁾. Fine pleomorphic microcalcifications vary in size and shape, are usually smaller than 0.5 mm in diameter, and are considered as high malignancy probability, as well as the fine linear branching microcalcifications^(3,4).

Punctate calcifications (smaller than 0.5 mm) have been associated to less than 2% malignancy, and can be classified as probably benign, depending on their distribution. Fine linear or fine linear branching calcifications are considered as highly suspicious, particularly when in segmental or linear distribution⁽³⁾, being associated with malignant lesions in 81% to 92% of cases. According to Liberman et al.⁽⁴⁾, approximately 41% of fine pleomorphic calcifications are associated to malignancy. Amorphous microcalcifications, in this BI-RADS edition indicated as a morphology of intermediate suspicion degree, presented a malignancy rate between 20% and 26%, specially associated with the segmental and linear distribution^(5,6).

It was therefore necessary to characterize microcalcifications according to their morphology, taking into account their distribution, and then classify them into BI-RADS categories. It is possible to observe that three subdivisions were suggested for category 4, with likely subjectivity in the choice between categories 4A, 4B and 4C, as there are two microcalcification groups with suspicious morphologies: those with intermediate suspicion (amorphous or indistinct and coarse heterogeneous) and those with a high probability of malignancy (fine pleomorphic and fine linear or fine linear branching).

Based on the knowledge of the predictive values of the different categories, the BI-RADS system determines that management recommendations should be suggested^(4,7).

The current recommendations advocate A PPV between 25% and 40% for breast cancer considering the lesions that are referred for biopsy⁽⁸⁾. The results of mammography sensitivity measurements range

from 68% to 88%^(9,10). According to Kerlikowski et al., sensitivity achieved 98% in fat containing breasts, decreasing to 63% in extremely dense breasts⁽¹⁰⁾. In the study developed by Kolb et al., mammography accuracy was 98.6%⁽⁸⁾.

It is a known fact that the accuracy of breast imaging studies may be affected by a number of factors, such as technical aspects, differences related to the characteristics of the population under study, patient's age, radiologist experience, use of double-reading technique or computer-aided detection systems – CADs), as well as the variability in the interpretation by the radiologist utilizing the BI-RADS⁽¹¹⁻¹³⁾.

Objective

The objective of the present study is to evaluate the accuracy of the BI-RADS classification in mammography, more specifically in what concerns the differentiation of benign lesions from malignant masses, description of frequency of the different mammographic findings, and evaluation of interobserver agreement.

MATERIALS AND METHODS

Imaging studies of 115 patients referred to a clinic in the Northwestern region of the Rio Grande do Sul State for being submitted to core biopsy, with previous mammographic diagnoses classified into BI-RADS categories 3, 4 or 5, were independently and blindly reviewed by two specialists in breast imaging diagnosis, both of them with more than ten years of experience, radiology specialist titles and/or qualification in mammography by the Colégio Brasileiro de Radiologia e Diagnóstico por Imagem. Terms, evaluation and recommendations were based on the BI-RADS, and the last version of the lexicon was utilized. The reviewed images were later compared with the anatomopathological results. The BI-RADS accuracy in the classification of mammograms was evaluated by means of sensitivity, specificity rates, (PPV) and negative predictive value (NPV) calculations for each one of the described characteristics, in the differentiation between benign and malignant lesions. Histological findings were utilized as a standard criterion.

Interobserver agreement for the final categories as a whole, and separately for each category, was calculated by means of the Cohen's kappa (κ) test, and the differences in the comparison groups were analyzed by means of the chi-squared test for category variables.

The observers described each lesion using the terminology from the fourth BI-RADS edition (Table 1) and the final mammography categorization included in the new BI-RADS subcategory 4 (Table 2). The radiologists did not receive any specific training on the use of BI-RADS, thus the criteria adopted by each of the radiologists were subjective, based on their previous knowledge of BI-RADS guidelines as well as on their individual experience. Once the lesions were duly described, they were classified as shown on Table 2.

Category 3 was included in the group of benign lesions and classes 4 (probably benign) and 5 were brought together as malignant. PPV and NPV were calculated for each class and description.

RESULTS

The present study population included 113 women and 2 men. The patients' ages ranged from 37 to 61 years with a mean age of 49 years (\pm 12 years).

Biopsies of 115 breast masses detected at mammography were performed. Sixty-seven of these lesions (58.3%) were benign and 48 (41.7%) were malignant.

Based on the BI-RADS for mammography, the cases were thus classified by the observer A as follows: 66 (57.4%) category 3, 30 (26.1%) category 4, and 19 (16.5%) category 5. The observer B classified the cases as follows: 36 (31.3%) category 3, 54 (47.0%) category 4 and 25 (21.7%) category 5. None of the cases were classified as categories 0, 1, 2 and 6.

For the observer A, the NPV was 76% and PPV was 51%. The sensitivity was 68%, specificity 76% and accuracy 75% (Table 3). For observer B, NPV was 83% and PPV 53%. Sensitivity was 87%, specificity 44% and accuracy 62% (Table 4).

Mammographic characteristics

The criteria described in the fourth BI-RADS edition were considered in the

Table 1 BI-RADS fourth edition terminology⁽²⁾.

Mammographic evaluation	Characteristics
<i>Calcifications – description</i>	
– Typically benign morphology	Coarse (or popcornlike) Cutaneous Vascular Lucent center Milk of calcium, needlelike Rodlike Rim or eggshell Suture Dystrophic Round/punctate (if alone)
– Morphology of intermediate suspiciousness	Amorphous Coarse heterogeneous
– Higher probability of malignancy	Fine pleomorphic Fine branching or casting
– Distribution	Grouped or clustered Linear Segmental Regional Diffuse or scattered
– Number	< 5 5–10 > 10
<i>Masses</i>	
– Shape	Round Oval Lobular Microlobulated (irregular)
– Margins	Circumscribed Microlobulated Indistinct or ill-defined Spiculated
– Density	Fat containing Low Equal High
<i>Architectural distortion</i>	
– Special cases	Intramammary lymph nodes Tubular density or dilated duct Global asymmetry Focal asymmetry
– Associated findings	Skin retraction Nipple retraction Trabecular thickening Skin lesion Axillary adenopathy

Table 2 BI-RADS fourth edition – final categories^(2,4).

Category	Definitions
1	Negative
0	Additional evaluation required
2	Benign findings
3	Probably benign findings
4A	Low malignancy suspicion
4B	Intermediate malignancy suspicion
4C	Moderate malignancy suspicion
5	Highly suggestive of malignancy

Table 3 Distribution of false-positive, true-positive, false-negative and true-negative results based on the pathological diagnosis and indicated diagnosis (observer A).

Test result	Positive disease		Negative disease		Total	
	n	(%)	n	(%)	n	(%)
T+ (classification 4, 5)	33	(68) TP	16	(33.3) FP	49	(100)
T- (classification 3)	15	(23.8) FN	51	(76.1) TN	66	(100)
Total	48	(41.7)	67	(58.3)	115	(100)
Parameters	Formula		%			
Sensitivity	TP/(TP + FN)		68 (with disease and positive test)			
Specificity	TN/(TN + FP)		76 (without disease and negative test)			
Positive predictive value	TP/(TP + FP)		51			
Negative predictive value	TN/(TN + FN)		76			
Accuracy	(TP + TN)/Total		75			

T+, positive test (lesion classified as BI-RADS 4 or 5); T-, negative test (lesion classified as BI-RADS 3). TP, true-positive; FP, false-positive; FN, false-negative; TN, true-negative.

Table 3 Distribution of false-positive, true-positive, false-negative and true-negative results based on the pathological diagnosis and indicated diagnosis (observer B).

Test result	Positive disease		Negative disease		Total	
	n	(%)	n	(%)	n	(%)
T+ (classification 4, 5)	42	(53.2) TP	37	(46.8) FP	79	(100)
T- (classification 3)	6	(16.7) FN	30	(83.3) TN	36	(100)
Total	48	(41.7)	67	(58.3)	115	(100)
Parameters	Formula		%			
Sensitivity	TP/(TP + FN)		87 (with disease and positive test)			
Specificity	TN/(TN + FP)		44 (without disease and negative test)			
Positive predictive value	TP/(TP + FP)		53			
Negative predictive value	TN/(TN + FN)		83			
Accuracy	(TP + TN)/Total		62			

T+, positive test (lesion classified as BI-RADS 4 or 5); T-, negative test (lesion classified as BI-RADS 3). TP, true-positive; FP, false-positive; FN, false-negative; TN, true-negative.

evaluation of masses and calcification demonstrated by mammographic images.

Evaluation of breasts density – The global agreement for the evaluation of breast density was moderate ($\kappa = 0.43$). The PPV for heterogeneously dense breasts was 43.8% for observer A and 39.6% for observer B.

Evaluation of lesions margins and shape – According to observer A, the round shape presented NPV of 75% and for the oval lesions, 71%. Lobular lesions presented a PPV of 70% and the microlobulated lesions presented a PPV of 90%. For observer B, the round shape presented a NPV of 70% and oval lesions, 66.7%. Lobular lesions presented a PPV of 75% and microlobulated lesions, 80%.

Evaluation of lesion margins – According to observer A, the NPV for circumscribed margins was 84.2%, while the PPVs for indistinct and spiculated margins

were 24.5% and 90%, respectively. According to observer B, the NPV for circumscribed margins was 80.2% and the PPVs for indistinct and spiculated margins were 25.4% and 83.3% respectively.

Evaluation of calcifications morphology – According to observer A, of the 76 reported calcifications, 23 (30%) were described as round, vascular or punctate, 6 (8%) as amorphous, 2 (2.6%) as coarse heterogeneous, 32 (42%) as fine branching and 12 (18%) as fine linear pleomorphic. The NPV for round calcifications was 56.5%. Amorphous calcifications presented a NPV of 66.6%. Calcifications described as fine branching presented a PPV of 72.7% and those described as fine linear pleomorphic presented a PPV of 91.6%. The two calcifications described as coarse heterogeneous were benign.

According to observer B, among the 68 reported calcifications, 40 (58.8%) were

described as round, vascular or punctate, 2 (2.9%) as amorphous, 5 (7%) as coarse heterogeneous, 4 (5.8%) as fine branching, and 17 (25%) as fine linear pleomorphic. The NPV for round calcifications was 65%. Of the two amorphous calcifications, one was benign and the other, malignant. Among the five calcifications described as coarse heterogeneous, two were malignant and three were benign, for a PPV of 40%. Those described as fine branching presented a PPV of 75% and for the fine linear pleomorphic ones the PPV was 94.7%.

Evaluation of calcifications distribution

– The calcifications were described as being grouped by observer A in 13 cases, with 8 cases classified as BI-RADS 4 and 5, with a PPV of 45%. Regional calcifications were identified in 12 cases, with 7 malignant and 5 benign, with a PPV of 58%. Scattered or diffuse calcifications presented a NPV of 42.8%. Segmental distribution was described in six cases, four of them malignant, with PPV of 66.6%. No case was described as linear ductal. Regional calcifications were described by observer B in 16 cases, with 9 being benign and 7 malignant, with a PPV of 45.7%. Calcifications were described as clustered in 20 cases, 14 of them included in BI-RADS categories 4 and 5, and 10 being malignant, with a PPV of 40%. Diffuse or scattered calcifications presented a NPV of 53.8%. Segmental distribution was described in nine cases, with five being malignant and four benign, with a PPV of 55%.

Architectural distortion – In the present study, the evaluation of architectural distortion (special cases and associated findings) could not be secondarily evaluated as the authors considered the number of presented cases as being insufficient.

Interobserver variability in mammography

The interobserver variability analysis for the description of mammographic lesions, using the Cohen's κ test, is shown on Table 5.

Evaluation of masses on mammography

There was a low global agreement ($\kappa = 0.40$) in the description of mass margins. Similarly, a low agreement rate was observed in the description of microlobulated

margins ($\kappa = 0.38$) and oval shaped masses ($\kappa = 0.32$).

A significant global agreement ($\kappa = 0.66$) was observed in the evaluation of mass margins, particularly for spiculated margins ($\kappa = 0.70$). Global agreement for masses density was moderate ($\kappa = 0.43$).

Mammographic evaluation of calcifications

Agreement was almost almost perfect as the presence of calcifications was evaluated ($\kappa = 0.88$). The observers demonstrated low global agreement in the description of calcifications morphology ($\kappa = 0.36$). The use of the terms “amorphous” and “fine branching” resulted in a moderate agreement ($\kappa = 0.41$ and $\kappa = 0.43$, respectively). Agreement was poor for the use of the terms “coarse heterogeneous” ($\kappa = 0.23$) and fine pleomorphic ($\kappa = 0.25$). Poor agreement was also observed in calcifications distribution evaluation ($\kappa = 0.24$) (Table 5). Agreement was also poor in the evaluation of presence of architectural distortion ($\kappa = 0.23$).

Interobserver agreement in relation to the presence of associated findings and special cases could not be secondarily evaluated, considering the low number of cases with architectural distortion.

Final categories evaluation

Poor agreement was observed in the evaluation of final categories.

The highest agreement rate was observed for lesions classified with high likelihood of malignancy, or category 5 ($\kappa = 0.42$). Poor agreement was observed for categories 3 ($\kappa = 0.30$), 4A ($\kappa = 0.15$), 4B ($\kappa = 0.13$) and 4C ($\kappa = 0.16$). For categories 4, even when grouped ($\kappa = 0.27$), poor agreement was observed.

For the final BI-RADS categories, poor interobserver agreement was observed ($\kappa = 0.32$) (Table 5).

DISCUSSION

In the present study, the utilization of criteria for breasts density, masses margins, shape, calcifications morphology and distribution was evaluated.

Mammography sensitivity ranged from 68% to 87% between the observers (iden-

Table 5 Interobserver variability in the description of mammographic lesions.

Description of lesions	Valor κ
<i>Masses</i>	
– Shape	0.40
– Margins	0.66
– Breast density	0.43
<i>Calcifications</i>	
– Morphology	0.36
– Distribution	0.24
BI-RADS for all categories	0.32

tification of malignant lesions in patients with breast cancer), and the NPV was high, ranging between 76% and 83% (identification of negative findings in cancer-regarding characteristics described by the BI-RADS. BI-RADS presented specificity between 76% and 44% (patients without the disease, with negative tests). The PPV (number of cancer cases for mammographic characteristics) ranged between 51% and 53% between the two observers, a rate not distant from results in the studies developed by Burnside et al.^(3,6) and Kerlikowske et al.⁽¹⁰⁾.

The mammographic accuracy ranged between 75% and 62% in the differentiation between benign and malignant lesions with the use of BI-RADS. The NPV for category 3 ranged between 76.1% and 83% between both observers, close to the values described by Roveda Junior et al.⁽¹⁴⁾

It is a known fact that there is a direct association between the increased mammographic density and an increase in the risk for development of breast cancer^(15,16). In the present study, the PPV for heterogeneously dense breasts were 43.8% for observer A and 39.6% for observer B. Moderate interobserver agreement ($\kappa = 0.43$) was observed in the evaluation of breasts density, differently from the findings described in the study developed by Nicholson et al.⁽¹³⁾, in which the interobserver agreement in the evaluation of breasts density was 78.4% for extremely dense breasts, and 51.2% for heterogeneously dense breasts, probably because of the different apparatuses utilized for images processing.

The present study suggests that the mass margins are useful in the prediction of malignancy, with a lower probability for carcinomas in lesions with well-defined

margins, and high probability in lesions with spiculated margins (non circumscribed), with NPV between 80% and 84% and PPV between 90% and 93%, respectively, for observers A and B, as described by Kestelman et al.⁽¹⁷⁾. It is known that, according to Nascimento et al.⁽¹⁸⁾, the sonographic method has also presented high PPV, at 82.4%, in the description of mass margins⁽¹⁹⁾.

As regards to round and oval shapes, these were associated to a high NPV, between 75% and 71% for observer A and between 70% and 66.7% for observer B. Microlobulated and lobular shapes presented a high PPV, between 90% and 70% for observer A, and 80% and 75% for observer B. In the present study a moderate interobserver agreement was observed in the global description of mass margins ($\kappa = 0.66$), in agreement with the findings reported by Kerlikowske et al.⁽¹⁰⁾.

In the present study, the observer A identified a high PPV in the description of fine branching microcalcifications and in fine linear pleomorphic microcalcifications (91,6%), a NPV of 56.5% for calcifications described as round, vascular or punctate and 66.6% for the amorphous ones. Observer B identified a PPV of 75% for fine branching calcifications, and 94.7% for fine linear pleomorphic calcifications, NPV of 65% for calcifications described as round, vascular or punctate, and 50% for the amorphous ones. For observer B, coarse heterogeneous microcalcifications presented a PPV of 40%. The present study is in agreement with the study developed by Melhado et al.⁽²⁰⁾, which demonstrated a progressive increase of PPV in BI-RADS categories 4A, 4B and 4C, suggesting that this subdivision contributes in a more precise manner in the indication of suspicious lesions.

However, poor interobserver agreement was observed for the description of calcifications morphology ($\kappa = 0.36$) and distribution at mammography ($\kappa = 0.24$) were low, as were those described in the literature according to Berg et al.⁽¹¹⁾ and Lazarus et al.⁽¹²⁾.

In the present study, the poor interobserver agreement in the evaluation of categories 4A ($\kappa = 0.15$), 4B ($\kappa = 0.13$), 4C ($\kappa = 0.16$) and combined categories 4 ($\kappa =$

0.27) were possibly associated to the high number of offered categories. A higher interobserver agreement was observed in category 5 ($\kappa = 0.42$).

CONCLUSION

The present study demonstrated that breast evaluation by mammography, utilizing the BI-RADS classification, is an accurate method in the differentiation between benign and malignant lesions. The most frequent findings related to neoplasias were masses with spiculated margins, microlobulated (irregular) shape, lobular mass, fine branching microcalcifications and linear fine pleomorphic calcifications. A high interobserver agreement was not achieved in the analysis of calcifications morphology and distribution, possibly because of the high number of offered categories. However, a progressive increase was observed in the PPVs in subcategories 4A, 4B and 4C, suggesting that such breakdown contributes in a more detailed manner for the identification of suspiciously malignant lesions. Such stratification may be useful for the communication of suspiciousness levels to physicians and patients, who may benefit from this information in their decision making processes.

It must also be highlighted that breast lesions related to BI-RADS category 3 presented a high NPV which should be considered as an relevant factor in the conservative management of such lesions with the purpose of avoiding unnecessary biopsies.

REFERENCES

1. American College of Radiology. Mammography. Illustrated Breast Imaging Reporting and Data System (BI-RADS). 4th ed. Reston: American College of Radiology; 2003.
2. Vizcaíno I, Gadea L, Andreo L. Short-term follow-up results in 795 nonpalpable probably benign lesions detected at screening mammography. *Radiology*. 2001;219:475–83.
3. Burnside ES, Ochsner JE, Fowler KJ, et al. Use of microcalcification descriptors in BI-RADS 4th edition to stratify risk of malignancy. *Radiology*. 2007;242:388–95.
4. Liberman L, Abramson AF, Squires CB, et al. The Breast Imaging Reporting and Data System: positive predictive value of mammographic features and final assessment categories. *AJR Am J Roentgenol*. 1998;171:35–40.
5. Berg WA, Arnoldus CL, Teferra E, et al. Biopsy of amorphous breast calcifications: pathologic outcome and yield at stereotactic biopsy. *Radiology*. 2001;221:495–503.
6. Burnside ES, Rubin DL, Fine JP, et al. Bayesian network to predict breast cancer risk of mammographic microcalcifications and reduce number of benign biopsy results: initial experience. *Radiology*. 2006;240:666–73.
7. Godinho ER, Koch HA. Submissão às recomendações do BI-RADS® por médicos e pacientes: análise preliminar de 3.000 exames realizados em uma clínica particular. *Radiol Bras*. 2004;37:21–3.
8. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology*. 2002;225:165–75.
9. Kerlikowske K, Smith-Bindman R, Ljung BM, et al. Evaluation of abnormal mammography results and palpable breast abnormalities. *Ann Intern Med*. 2003;139:274–84.
10. Kerlikowske K, Grady D, Barclay J, et al. Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System. *Nat Cancer Inst*. 1998;90:1801–9.
11. Berg WA, Campassi C, Langenberg P, et al. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. *AJR Am J Roentgenol*. 2000;174:1769–77.
12. Lazarus E, Mainiero MB, Schepps B, et al. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. *Radiology*. 2006;239:385–91.
13. Nicholson BT, LoRusso AP, Smolkin M, et al. Accuracy of assigned BI-RADS breast density category definitions. *Acad Radiol*. 2006;13:1143–9.
14. Roveda Junior D, Piato S, Oliveira VM, et al. Valores preditivos das categorias 3, 4 e 5 do sistema BI-RADS em lesões mamárias nodulares não-palpáveis avaliadas por mamografia, ultrasonografia e ressonância magnética. *Radiol Bras*. 2007;40:93–8.
15. Boyd NF, Dite GS, Stone J. Heritability of mammographic density, a risk factor for breast cancer. *N Engl J Med*. 2002;347:886–94.
16. Warner E, Lockwood G, Tritchler D, et al. The risk of breast cancer associated with mammographic parenchymal patterns: a meta-analysis of the published literature to examine the effect of method of classification. *Cancer Detect Prev*. 1992;16:67–72.
17. Kestelman FP, Souza GA, Thuler LC, et al. Breast Imaging Reporting and Data System – BI-RADS®: valor preditivo positivo das categorias 3, 4 e 5. Revisão sistemática de literatura. *Radiol Bras*. 2007;40:173–7.
18. Nascimento JHR, Silva VD, Maciel AC. Acurácia dos achados ultrassonográficos do câncer de mama: correlação da classificação BI-RADS® e achados histológicos. *Radiol Bras*. 2009;42:235–40.
19. Arantes Pereira FP. BI-RADS® ultrassonográfico: análise de resultados iniciais [editorial]. *Radiol Bras*. 2009;42(4):vii–viii.
20. Melhado VC, Alvares BR, Almeida OJ. Correlação radiológica e histológica de lesões mamárias não-palpáveis em pacientes submetidas a marcação pré-cirúrgica, utilizando-se o sistema BI-RADS. *Radiol Bras*. 2007;40:9–11.