BREAST IMAGING REPORTING AND DATA SYSTEM – BI-RADS®: POSITIVE PREDICTIVE VALUE OF CATEGORIES 3, 4 AND 5. A SYSTEMATIC LITERATURE REVIEW*

Fabíola Procaci Kestelman[1], Gustavo Antônio de Souza[2], Luiz Claudio Thuler[3], Gabriela Martins[4], Vivianne Aguilera Rolim de Freitas[4], Ellyete de Oliveira Canella[5]

Abstract **OBJECTIVE:** To review the literature about the positive predictive value of Breast Imaging Reporting and Data System (BI-RADS®) categories 3, 4 and 5. **MATERIALS AND METHODS:** A research was performed in the online Medline database, entering the terms "predictive value" and "BI-RADS". Eleven studies were included in this review. **RESULTS:** The positive predictive values ranged respectively between 0% and 8%, 4% and 62%, and 54% and 100% for BI-RADS® categories 3, 4 and 5. Three studies have also evaluated morphological criteria with higher positive predictive value in mammography, mass with spiculated margins being the finding with highest positive predictive value for malignancy. **CONCLUSION:** A high variability was found in the reviewed literature among positive predictive values of BI-RADS® categories 3, 4 and 5, although methodological differences have been identified, limiting the comparative analysis.
Keywords: Mammography; BI-RADS®; Breast cancer.

Resumo *Breast Imaging Reporting and Data System – BI-RADS®: valor preditivo positivo das categorias 3, 4 e 5. Revisão sistemática da literatura.*
OBJETIVO: Avaliar artigos, na literatura, que verificam o valor preditivo positivo das categorias 3, 4 e 5 do Breast Imaging Reporting and Data System (BI-RADS®). **MATERIAIS E MÉTODOS:** Foi realizada pesquisa na base de dados Medline utilizando os termos "predictive value" e "BI-RADS". Foram incluídos 11 artigos nesta revisão. **RESULTADOS:** O valor preditivo positivo das categorias 3, 4 e 5 variou entre 0% e 8%, 4% e 62%, 54% e 100%, respectivamente. Três artigos avaliaram, concomitantemente, os critérios morfológicos das lesões que apresentaram maior valor preditivo positivo na mamografia, sendo nódulo espiculado o critério com maior valor preditivo positivo. **CONCLUSÃO:** Houve grande variabilidade do valor preditivo positivo das categorias 3, 4 e 5 do BI-RADS® em todos os estudos, porém foram identificadas diferenças metodológicas que limitaram a comparação desses estudos.
Unitermos: Mamografia; BI-RADS® ; Câncer de mama.

INTRODUCTION

Breast cancer is the main cause of death due to neoplasm amongst women in Brazil, corresponding to the second most in-

cident type of cancer, according to Instituto Nacional de Câncer statistics[1]. The decrease in mortality rate depends on an early detection, an adequate therapeutic planning, and the adoption of the annual mammography as screening method[2].

Mammography presents high sensitivity in the detection of clinically occult breast cancer. A review of clinical trials evaluating the performance of the method has demonstrated that the sensitivity ranged between 71% and 98% for the annual screening mammography[3]. However, many lesions deemed suspicious with indication for histopathological evaluation correspond to benign alterations. In the United States, the positive predictive value (PPV) of biopsies performed because of mammographic findings, that is to say, the total of diagnosed malignant lesions in relation to the total number of biopsies performed, ranges between 15% and 40%[4–6].

The cost and morbidity of the procedures for the diagnosis of these lesions are taken into consideration to confirm the adoption of mammography as a screening method[7].

One of the difficulties in the evaluation of mammograms is that the greatest part of the diagnosed lesions does not present pathognomonic characteristics. Knutzen & Gisvold[8] have studied the probability of malignancy in several categories of non-palpable lesions detected by mammography, and observed that, if the morphological criteria of these lesions were taken into consideration, the rate of malignant lesions among women submitted to biopsy could reach 40%.

Aiming at reducing the discordance in the interpretation of mammographic findings and standardizing the terms for characterization and reporting, the American College of Radiology published, in 1993, the Breast Imaging Reporting and Data

System (BI-RADS®)[9]. New editions were released in 1995, 1998 and 2003[10–12]. According to the latest (fourth) edition of BI-RADS®[12], the studies are classified with basis on the lesions' suspicion level into: category 1 – without positive findings; category 2 – benign findings; category 3 – probably benign findings; category 4 – suspicious findings; category 5 – findings highly suggestive of malignancy. Lesions requiring additional evaluation, for example with ultrasound, are classified into category 0, and those with previously confirmed histopathological diagnosis of malignancy, into category 6.

The BI-RADS® latest edition included suggestions for evaluation of breast lesions with ultrasound and magnetic resonance imaging. Notwithstanding its great utility already observed in relation to the BI-RADS® for mammography, the use of this lexicon is still recent, generating some criticism and considerations[13,14].

Some studies evaluate the capacity to predict malignancy of categories 3, 4 and 5 where lesions presenting some level of suspicion would be included. A way to evaluate the performance of each BI-RADS® category is to analyse the results of lesions submitted to biopsy, and calculate the PPV considering the histopathological result as the gold-standard.

This literature review is aimed at evaluating the PPV of BI-RADS® categories 3, 4 and 5. Additionally, morphological criteria utilized for stratification of lesions were analyzed.

MATERIALS AND METHODS

The literature review was performed by means of a search in the online Medline database, by entering the terms "predictive value" and "BI-RADS®". The publications included in the present review have met the following criteria: original articles evaluating the PPV of BI-RADS® categories 3, 4 and 5; PPV analysis based on histopathological results; studies published from 1998; papers written in Portuguese, English or French. Studies evaluating only one or two categories, or whose online abstract was not available were excluded.

The selected original articles were integrally evaluated in order to confirm their compliance with the above mentioned criteria. Nine of the 33 studies found in the search have met these criteria. By means of the bibliographic references in these nine studies, we have found other two compliant studies, totalling 11 articles.

The data extracted from the articles were registered on an Excel® (Microsoft) worksheet. The following items were included: study period; number of lesions evaluated; patients' ages; type of lesion (palpable/non-palpable); biopsy method; clinical, radiological or surgical management of the lesions submitted to percutaneous biopsy; PPV of BI-RADS® categories 3, 4 and 5; morphological criteria analysis.

The categories' PPVs have been extracted from the articles or calculated through the data available, considering the number of lesions in one of the categories with malignant diagnosis divided by the total number of lesions classified into this category, and finally multiplied by 100. In this case, the histological results classified as malignant were ductal carcinoma *in situ* or any other primary, invasive breast tumor, according to BI-RADS® recommendations[12].

Only one study has evaluated both lesions submitted to some biopsy method and lesions under radiological follow-up. In this case, the PPV was calculated considering only the lesions initially submitted to biopsy[15], with the histopathological criteria as the gold-standard.

RESULTS

The 11 articles reviewed reported the data required for determination of categories 3, 4 and 5 PPV. Data from these studies are summarized in Table 1.

In three articles, information was obtained exclusively from patients submitted to surgical biopsy[16–18]. Five studies reported the assessment of patients submitted to percutaneous biopsy[19–23], and two studies evaluated results of patients submitted to percutaneous biopsy and patients submitted to surgical biopsy in conjunction[15,24]. Liberman et al. have studied separately patients submitted to percutaneous biopsy and patients submitted to surgical biopsies[25].

Table 1 Studies on breast lesions with PPV for BI-RADS® categories 3, 4 and 5.

Study	Period	N	Age (years)	Physical examination	Biopsy method	Carcinomas prevalence	PPV according to BI-RADS® category		
							BI-RADS® 3	BI-RADS® 4	BI-RADS® 5
Liberman et al.[25], 1998	Jan./1996–Nov./1996	492	27–92	Non-palpable	SB	46%	0% (0/8)	34% (120/355)	81% (105/129)
Bérubé et al.[19], 1998	Jul./1994–Jun./1995	193	?	?	PB (CB and MT)	37%	0% (0/2)	18% (25/141)	92% (46/50)
Orel et al.[16], 1999	Apr./1991–Dec./1996	495	22–89	?	PB (CB)	12%	0% (0/16)	4% (17/397)	54% (165/170)
Lacquement et al.[24], 1999	Jan./1995–Jun./1997	1,312	?	Non-palpable	SB	34%	2% (3/141)	30% (279/436)	97% (165/170)
Tate et al.[20], 2001	Feb./1994–Dec./1999	68	31–88	Palpable and non-palpable	SB and PB (FNAB, CB and MT)	23%	3% (9/322)	23% (54/234)	92% (97/106)
Margolin et al.[21], 2001	Jan./1994–Dec./1998	1,080	?	Palpable and non-palpable	PB (CB, MT and ABBI)	18%	2% (2/87)	17% (159/961)	90% (36/40)
Ball et al.[17], 2002	Jan./1994–Jun./1999	1,333	37–90	Non-palpable	SB	11%	3% (1/34)	10% (100/967)	92% (24/26)
Travade et al.[22], 2002	Oct./1999–Jan./2001	109	29–80	Non-palpable	PB (MT)	38%	0% (0/10)	26% (18/68)	77% (24/71)
Mendez et al.[23], 2003	Aug./2000–Dec./2002	252	25–89	?	PB (MT)	28%	0% (0/12)	20% (41/206)	88% (30/34)
Tan et al.[18], 2004	Jan./1995–Dec./2000	947	31–85	Palpable and non-palpable	SB	16%	4% (7/156)	15% (116/757)	79% (27/34)
Zonderland et al.[15], 2004	Jun./2000–Jun./2001	497	21–92	Palpable and non-palpable	SB and PB (FNAB, CB)	16%	5% (10/201)	27% (49/180)	84% (16/19)
		210				54%	8% (4/49)	62% (39/62)	100% (57/57)

SB, surgical biopsy; PB, percutaneous biopsy; CB, core biopsy; MT, mammotomy; FNAB, fine needle aspiration biopsy; ABBI, advanced breast biopsy instrumentation.

As regards the studies based on percutaneous biopsies, four of them refer to the follow-up of patients both by radiological follow-up of lesions with benign histological diagnosis, and by analysis of histological results from later surgical biopsies, for example, in cases with initial histological diagnosis of atypical ductal hyperplasia[20-22,25]. In the other studies, these data are incomplete. Bérubé et al.[19] have reported a new biopsy in cases where there was disagreement between the histological report and the radiological aspect of the lesion. Two studies have correlated histological results of atypical hyperplasia with those of surgical biopsy[23,24]. Zonderland et al.[15] describe the clinical management of lesions by means of data from de institution itself and from a national register of histopathological studies results.

As regards selection of patients considering the physical examination of their breasts, the studies have not followed a homogeneous criterion. Eight of them report physical examination, with four evaluating only non-palpable lesions[16,17,22,25] and the others have report assessment of both palpable and non-palpable lesions[15,18,21,24]. Three articles have not clearly described the type of lesions studied[19,20,23], however, they have assessed patients submitted to ultrasound- or stereotactic-guided percutaneous biopsy, which may indicate that these lesion are clinically occult. In the study where Liberman et al.[25] have evaluated patients submitted to percutaneous biopsy, these data are not sufficiently defined, but the lesions have also been submitted to ultrasound- or stereotactic-guided biopsy.

PPV of categories 3, 4 and 5

In the 11 studies, the PPV ranged between 0% and 8% for category 3 (median, 2%), 4% and 63% for category 4 (median, 21%), and between 54% and 100% for category 5 (median, 89%)(Figure 1). In the four studies where the method for obtaining the histopathological result was exclusively surgical biopsy, the PPV found for categories 3, 4 and 5 ranged, respectively, between 0% and 5%, 26% and 34%, and between 81% and 97%[16–18,25]. In the six studies evaluating lesions submitted only to percutaneous biopsy, the PPV for cat-
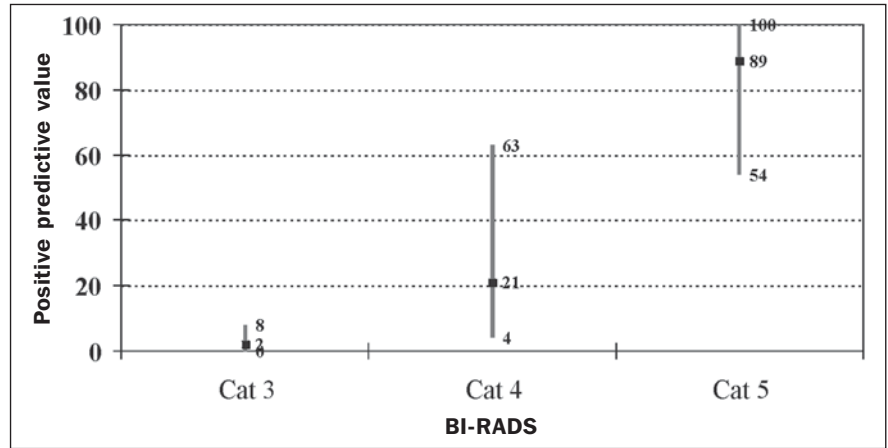

Figure 1. Positive predictive values reported in studies for BI-RADS® categories.

egories 3, 4 and 5 ranged, respectively, between 0% and 4%, 4% and 20%, and between 54% and 92%[19–23,25].

The four studies evaluating only non-palpable lesions[16,17,22,25] presented a PPV variation, for category 3, between 0% and 2%, for category 4, between 20% and 34%, and, for category 5, between 77% and 97%. The four articles reporting results from palpable and non-palpable lesions[15,18,21,24] demonstrated a PPV variation, for category 3, between 3% and 8%, for category 4, between 10% and 63%, and for category 5, between 84% and 100%.

Morphological criteria with higher PPV

Three articles have mentioned the morphological characteristics of the lesions with major association with malignancy. Liberman et al.[25] and Bérubé et al.[19] have evaluated morphological characteristics of lesions as proposed by the second edition of BI-RADS®[12], including, for analysis of masses, the evaluation of margins and shape, and for calcifications, the morphology and distribution (Table 2). Mendez et

al.[23] have classified the lesions into microcalcifications, asymmetrical density, circumscribed mass, spiculated or with microcalcifications, and asymmetrical density with microcalcifications. In the three studies, spiculated mass has been the lesion with highest PPV[19,23,25]. As regards masses, Liberman et al.[25] have observed that the criteria with higher PPV were spiculated margins and irregular shape; and, as regards microcalcifications, were linear morphology and segmental or linear distribution. Bérubé et al.[19] have reported that the morphological criteria with higher PPV were dense masses, massed with spiculated margins, and linear microcalcifications.

DISCUSSION

The studies analyzed with the purpose of evaluating the PPV for BI-RADS® categories 3, 4 and 5 demonstrated a significant difference in the cancer detection among these categories. However, the methodology utilized in those studies was

Table 2 Radiological findings with highest PPV, according to BI-RADS®.

Study	Morphological criteria	PPV
Liberman et al.[25], 1998	Mass with spiculated margin	81%
	Irregularly shaped mass	73%
	Microcalcifications with linear morphology	81%
	Microcacifications with segmental distribution	74%
	Microcalcifications with linear distribution	68%
Bérubé et al.[19], 1998	Dense mass	90%
	Mass with spiculated margin	91%
	Microcalcifications with linear morphology	80%

quite heterogeneous, limiting the comparison between results. Some studies presented limited information regarding patients selection. Some of them do not mention if the lesions were palpable or nonpalpable. In the articles where there is no reference to the type of the studied lesions, the histopathological result was obtained by means of ultrasound- or stereotactic-guided percutaneous biopsy, suggesting that the lesions were clinically occult. The variation of the carcinomas prevalence in these studies and, consequently, of PPV for categories, may be related with the differences in the selection of patients for surgical or percutaneous biopsy.

Another factor that may influence the carcinomas prevalence is the radiological follow-up of lesions with benign histopathological diagnosis. In studies with follow-up of patients after core needle (14 gauge) biopsy, the frequency of non-diagnosed carcinomas ranged between 0.3% and 8.2%, and, with the radiological follow-up, 70% of these carcinomas were immediate false negative, and 30%, delayed false negative[26,27].

Additionally, as regards the number of carcinomas diagnosed by percutaneous biopsy, another variable should be considered: the management of underestimated lesions. For example, a lesion diagnosed as atypical ductal hyperplasia, after percutaneous biopsy, may correspond to a ductal carcinoma in situ or an invasive ductal carcinoma. As regards lesions diagnosed as atypical ductal hyperplasia, when submitted to core needle (14 gauge) biopsy, 20% to 56% correspond to carcinomas in surgical biopsies, while 0% to 38% of lesions submitted to directional vacuum-assisted biopsy (mammotomy) with 14- or 11-gauge needle are underestimated as atypical ductal hyperplasia. These data demonstrate that, when the volume of biopsy tissue acquired is higher, the underestimation rate is lower, although not non-existent[28].

Considering these data, the studies analyzed also demonstrate heterogeneous characteristics. Six[15,20,22–25] of the eight articles[15,19–25] where the histopathological results have been obtained from lesions submitted to percutaneous biopsy report the radiological follow-up of benign lesions and surgical biopsy of determined

benign lesions, for example, atypical ductal hyperplasia. Considering that these two variables affect the number of malignant diagnoses, studies presenting homogeneity between these aspects may show differences in their results.

Another important factor for the analysis of the categories' PPV is related to the use of BI-RADS®, which may be limited by deficiencies in the classification or training of radiologists[29,30]. Berg et al.[31] have evaluated inter- and intraobserver variability in the utilization of the BI-RADS® terminology. Five mammography-experienced radiologists have assessed 103 routine mammograms. The rate of agreement among the radiologists (kappa statistical method) ranged between 0.16 and 0.77 for the several mammographic findings, showing moderate variability, and 0.37 for BI-RADS® categories, corresponding to poor agreement. Orel et al.[16], analyzing the mammographic findings in different institutions, have observed that some patients referred for biopsy presented benign lesions classified as BI-RADS® category 2 and, therefore, biopsy would not be indicated. This discrepancy was associated with the interobserver variation, both for describing the lesions and for recommending biopsy, and also was associated with variations in the experience of the radiologists involved in the mammograms analysis. Bérubé et al.[19] have associated the low PPV obtained for category 4 with the fact of the BI-RADS® nomenclature being descriptive and poorly specific.

Contrarily to Bérubé et al.[19], Zonderland et al.[15] have found higher PPVs for all the categories, in comparison with results found in American studies. This finding has been attributed to the fact that, in the American studies, there is a tendency to get a higher number of positive diagnoses based on biopsy in order to reduce the number of false-negative results from mammography. This affects the selection of patients for histopathological investigation. The same finding is reported by a study comparing routine mammograms performed in the United States and in the United Kingdom[32].

As regards PPVs of categories, the BI-RADS® suggests values below 2% for category 3 and above 95% for category 5,

while five studies have found values above the one suggested for category 3[15,18,21,23,24], and in nine articles the values were lower than expected for category 5[17–25].

Finally, the present review demonstrates that in all of the studies there was a high variability in PPVs of BI-RADS® categories 3, 4 and 5. The comparison among results from the 11 studies is limited by the heterogeneity in the criteria for patients selection, biopsy methods, and, in the case of the studies with results from percutaneous biopsy, by the heterogeneity in the radiological follow-up of both benign and underestimated lesions. However, there is a scale for malignancy prediction allowing the definition, with a relative assurance, of patients with higher risk of breast cancer. Additionally, the studies demonstrate that the presence of spicutated mass presents a high association with malignancy.

Acknowledgments

REFERENCES

1. INCA. Instituto Nacional do Câncer, Ministério da Saúde. Estimativa da incidência e da mortalidade por câncer no Brasil. Acessado em: 22/9/2004. Disponível em: http://www.inca.gov.br/estimativas/2003/index.asp?link=conteudo_view.asp&ID=4

2. Newman LA, Sabel M. Advances in breast cancer detection and management. Med Clin North Am 2003;87:997–1028.

3. Thuler LC. Considerações sobre a prevenção do câncer de mama feminino. Rev Bras Cancerol 2003;49:227–238.

4. Ciatto S, Cataliotti L, Distante V. Nonpalpable lesions detected with mammography: review of 512 consecutive cases. Radiology 1987;165:99–102.

5. Cyrlak D. Induced costs of low-cost screening mammography. Radiology 1988;168:661–663.

6. Hall FM, Storella JM, Siverstone DZ, Wyshak G. Nonpalpable breast lesions: recommendation for biopsy based on suspicion of carcinoma at mammography. Radiology 1988;167:353–358.

7. Heywang-Köbrunner SH, Dershaw DD, Schreer I. Diagnostic breast imaging: mammography, sonography, magnetic resonance imaging, and in-

terventional procedures. 2nd ed. New York, NY: Georg Thieme Verlag, 2001;393.

8. Knutzen AM, Gisvold JJ. Likelihood of malignant disease for various categories of mammographically detected, nonpalpable breast lesions. Mayo Clin Proc 1993;68:454–460.

9. American College of Radiology. Breast Imaging Reporting and Data System (BI-RADS). Reston, VA: American College of Radiology, 1993.

10. American College of Radiology. Breast Imaging Reporting and Data System (BI-RADS). 2nd ed. Reston, VA: American College of Radiology, 1995.

11. American College of Radiology. Breast Imaging Reporting and Data System (BI-RADS®). 3rd ed. Reston, VA: American College of Radiology, 1998.

12. American College of Radiology. Breast Imaging Reporting and Data System (BI-RADS®). 4th ed. Reston, VA: American College of Radiology, 2003.

13. Chala LF, Barros N. ACR BI-RADS™ na ultra-sonografia. Radiol Bras 2004;37(2):III–IV.

14. Camargo Júnior HSA. BI-RADS®-ultra-som: vantagens e desvantagens dessa nova ferramenta de trabalho. Radiol Bras 2005;38:301–303.

15. Zonderland HM, Pope TL Jr, Nieborg AJ. The positive predictive value of the Breast Imaging Reporting and Data System (BI-RADS) as a method of quality assessment in breast imaging in a hospital population. Eur Radiol 2004;14: 1743–1750.

16. Orel SG, Kay N, Reynolds C, Sullivan DC. BI-RADS categorization as a predictor of malignancy. Radiology 1999;211:845–850.

17. Ball CG, Butchart M, MacFarlane JK. Effect on biopsy technique of the breast imaging reporting and data system (BI-RADS) for nonpalpable mammographic abnormalities. Can J Surg 2002; 45:259–263.

18. Tan YY, Wee SB, Tan MP, Chong BK. Positive predictive value of BI-RADS categorization in an Asian population. Asian J Surg 2004;27:186–191.

19. Bérubé M, Curpen B, Ugolini P, Lalonde L, Oui-met-Oliva D. Level of suspicion of a mammographic lesion: use of features defined by BI-RADS lexicon and correlation with large-core breast biopsy. Can Assoc Radiol J 1998;49:223––228.

20. Tate PS, Rogers EL, McGee EM, et al. Stereotactic breast biopsy: a six-year surgical experience. J Ky Med Assoc 2001;99:98–103.

21. Margolin FR, Leung JW, Jacobs RP, Denny SR. Percutaneous imaging-guided core breast biopsy: 5 years' experience in a community hospital. AJR Am J Roentgenol 2001;177:559–564.

22. Travade A, Isnard A, Bagard C, et al. Macrobiopsies stéréotaxiques par système à aspiration 11-G: à propos de 249 patientes. J Radiol 2002;83(9 Pt 1):1063–1071.

23. Mendez A, Cabanillas F, Echenique M, Malekshamran K, Perez I, Ramos E. Mammographic features and correlation with biopsy findings using 11-gauge stereotactic vacuum-assisted breast biopsy (SVABB). Ann Oncol 2004;15:450–454.

24. Lacquement MA, Mitchell D, Hollingsworth AB. Positive predictive value of the Breast Imaging Reporting and Data System. J Am Coll Surg 1999;189:34–40.

25. Liberman L, Abramson AF, Squires FB, Glass-

man JR, Morris EA, Dershaw DD. 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.

26. Jackman RJ, Nowels KW, Rodriguez-Soto J, Marzoni FA Jr, Finkelstein SI, Shepard MJ. Stereotactic, automated, large-core needle biopsy of nonpalpable breast lesions: false-negative and histologic underestimation rates after long-term follow-up. Radiology 1999;210:799–805.

27. Lee CH, Philpotts LE, Horvath LJ, Tocino I. Follow-up of breast lesions diagnosed as benign with stereotactic core-needle biopsy: frequency of mammographic change and false-negative rate. Radiology 1999;212:189–194.

28. Liberman L. Percutaneous image-guided core breast biopsy. Radiol Clin North Am 2002;40: 483–500.

29. Liberman L, Menell JH. Breast Imaging Reporting and Data System (BI-RADS). Radiol Clin North Am 2002;40:409–430.

30. Godinho ER, Koch HA. Breast Imaging Reporting and Data System (BI-RADS™): como tem sido utilizado? Radiol Bras 2004;37:413–417.

31. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. AJR Am J Roentgenol 2000;174:1769–1777.

32. Smith-Bindman R, Chu PW, Miglioretti DL, et al. Comparison of screening mammography in the United States and the United Kingdom. JAMA 2003;290:2129–2137.