

Quality of the interpretation of diagnostic mammographic images*

Qualidade da interpretação do diagnóstico mamográfico

Hilton Koch¹, Melissa Vieira Koch e Castro²

Abstract **OBJECTIVE:** To demonstrate the knowledge of mammogram readers working in the public healthcare system in the State of Rio de Janeiro, RJ, Brazil, and to evaluate their progress in the early diagnosis of breast cancer after a training course specifically developed for medical professionals. **MATERIALS AND METHODS:** A group of 53 physicians with experience in mammography reports were invited. A pre-test was given to assess their initial knowledge level. Afterwards, they were trained by experts mammographers, and for final conclusion, requested to take a post-test for comparison and evaluation of gained knowledge. **RESULTS:** The course, with emphasis on theoretical classes, has not resulted in a significant improvement on the quality of mammogram reading, highlighting the persistence of errors in morphological description of fundamental lesions of the breast, in the classification of such lesions according to the BI-RADS[®], besides the lack of coherence between the BI-RADS classification and follow-up recommendation as observed in both the pre- and post-test. **CONCLUSION:** The authors conclude that the mammogram readers have demonstrated insufficient knowledge in relation to early imaging diagnosis of breast cancer, and that the theoretical training has not resulted in a significant improvement on the quality of mammogram reading.
Keywords: Quality; Mammographic diagnosis; Technical parameters.

Resumo **OBJETIVO:** Demonstrar o conhecimento mamográfico dos médicos interpretadores que trabalham na rede de saúde pública do Estado do Rio de Janeiro e avaliar o conhecimento adquirido após um curso elaborado com o objetivo de capacitar profissionais médicos no diagnóstico precoce do câncer de mama. **MATERIAIS E MÉTODOS:** Foram convidados 53 médicos que laudam exames mamográficos para o treinamento. Esses médicos eram submetidos a um pré-teste, no qual se avaliava o grau de conhecimento inicial. Depois, foram lecionadas aulas previamente elaboradas por mamografistas experientes, e para conclusão do curso esses médicos eram submetidos a um pós-teste para avaliação do conhecimento adquirido. **RESULTADOS:** O curso de capacitação de profissionais médicos, com ênfase em aulas teóricas, não mostrou aumento significativo na qualidade da interpretação mamográfica, destacando-se a persistência do erro na descrição morfológica das lesões fundamentais da mama, erro da classificação pelo sistema de padronização das lesões mamárias (BI-RADS[®]), falta de coerência entre a classificação BI-RADS adotada e a recomendação de conduta, tanto no pré-teste como no pós-teste. **CONCLUSÃO:** Concluiu-se que os médicos interpretadores mostram conhecimento insuficiente em relação ao diagnóstico precoce por imagem do câncer de mama e que o curso teórico não mostrou aumento significativo na qualidade da interpretação mamográfica.
Unitermos: Qualidade; Diagnóstico mamográfico; Parâmetros técnicos.

Koch HA, Castro MVK. Quality of the interpretation of diagnostic mammographic images. *Radiol Bras.* 2010;43(2):97–101.

INTRODUCTION

Breast radiology has been undergoing changes over the last years. The adoption

of mammography as a method for screening breast diseases in asymptomatic patients has intensified the mammographic diagnosis complexity and has highlighted the need to evaluate the different factors that may influence the variability of such diagnosis, such as the definition of technical parameters and the level of learning of the observer interpreting the images^(1,2). In 1991, Colégio Brasileiro de Radiologia (CBR) launched its Mammography Quality Control Program, offering radiologists the conditions to improve the quality of their studies, which reached a peak in 1997, when 75% of the mammography equip-

ment and accessories had been certified by CBR^(3,4). As regards the mammographic interpretation, specialized and training courses in mammography may enhance the effectiveness of the interpreting physician⁽⁵⁾. The CBR has approved the standardization of mammographic reports in accordance with the standards set by the Breast Imaging Reporting and Data System – BI-RADS^{®(6)} to improve diagnosis; however, it was observed that only 15% of the mammography centers applied the BI-RADS in their reports⁽⁷⁾. It is interesting to observe that even with unsatisfactory data, the courses organized throughout the coun-

* Study developed at Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, RJ, Brazil.

1. PhD, Titular Professor and Head of the Department of Radiology at Faculdade de Medicina da Universidade Federal do Rio de Janeiro (UFRJ), Head of the Division of Radiology at Santa Casa da Misericórdia do Rio de Janeiro, Rio de Janeiro, RJ, Brazil.

2. PhD, Responsible by the Unit of Breast Diagnosis at Division of Radiology at Santa Casa da Misericórdia do Rio de Janeiro, Rio de Janeiro, RJ, Brazil.

Mailing address: Dr. Hilton Koch. Estrada de Itajuru, 425, Itanhangá, Barra da Tijuca. Rio de Janeiro, RJ, Brazil, 22641-190. E-mail: hakoch@uol.com.br

Received July 17, 2009. Accepted after revision February 26, 2010.

try emphasize innovative methods, as for example, the use of magnetic resonance imaging, without highlighting the importance of baseline diagnosis achieved by means of mammography. Therefore, there is a gap between what the interpreting physicians need to know to perform an early diagnosis of breast disease and what they really know.

Objective

The main objectives of the present study were the evaluation of the knowledge on mammography possessed by practitioners reading such images, and later, the assessment of their level of learning after undergoing a specific course on breast studies by means of mammography, with the objective of improving the breast diseases diagnosis.

MATERIALS AND METHODS

The mentioned course was initially idealized in a consensus meeting attended by experienced specialized radiologists and mastologists, public organs officers and representatives of the professional council. In this meeting, the course curriculum and schedule were defined. The schedule comprised nine classes basically aimed at the

learning on BI-RADS classification. Classes were to be one hour long at most, and after class the students would have time as long as needed to solve their doubts. On a later step, the classes were standardized and a manual was developed.

A pre-course test was applied in order to assess the attendees' degree of knowledge and, upon the course conclusion, another test was applied to evaluate their level of learning through a review of the same cases presented on the first test along with new cases.

The mammographic studies presented on both tests were collected at the Unit of Breast Diagnosis – Division of Radiology of Santa Casa da Misericórdia do Rio de Janeiro, RJ, Brazil, selected from the period comprised between September 20, 2005 and December 15, 2005. The images were digitized and saved by means of a computer software.

The cases were presented by means of a datashow device and attendees were asked to fill out a form that followed the standardized BI-RADS system. Such forms were compared in order to assess the level of learning.

The selected physicians participating in this project were radiologists and masto-

logists responsible for the preparation of mammographic reports for centers of the public health system (Sistema Único de Saúde – SUS) in the State of Rio de Janeiro, Brazil.

Five courses were given, four of them for physicians from the SUS network and one for residents and trainees, in a total of 53 evaluated participants, as follows: course 1 (Avon) with 10 physicians; course 2 (Avon) with 12 physicians; course 3 (residents and trainees) with 17 physicians; course 4 (Avon 3) with 7 physicians; and course 5 (Avon 4) with 7 physicians. However, other physicians were not submitted to the post-course test, as some attended only part of the course, and others could not make the test for different reasons, so that for these cases, it was not possible to compare the results.

The selected cases covered breasts with no radiological alteration, operated breasts (with silicone implants and mastoplasty), breasts with benign calcifications, breasts with typically malignant masses, breasts with typically benign masses, breasts with architectural distortions, besides studies with technical errors (processing) (Table 1). *Note:* Suspicious microcalcifications were not evaluated in the tests, because of the

Table 1 Cases valuated in the pre- and post-course tests.

No.	Form/case No.	Radiological finding	BI-RADS category	Presented on course
1 Pre-course test	3167805	Regular, well-defined mass in the left breast	0	1, 2, 3, 4, 5
2 Pre-course test	3177605	Architectural distortion in the upper outer quadrant of the left breast	4C	1, 2, 3, 4, 5
3 Pre-course test	3252005	Benign round calcifications, some of them with lucent center	2	1, 2, 3, 4, 5
4 Pre-course test	3256005	Retroglandular prosthesis	2	1, 2, 3, 4, 5
5 Pre-course test	32688805	Architectural distortion associated with microcalcifications and skin thickening on the left breast	5	1, 2, 3, 4, 5
6 Pre-course test	3283405	Vascular calcifications	2	1, 2, 3, 4, 5
7 Pre-course test	3329205	Predominantly fatty breasts (roll marks)	1	1, 2, 3, 4, 5
8 Post-course test	3407905	Reduction mastoplasty in fatty breast	2	1, 3, 4, 5
9 Post-course test	3401905	Mass in the upper outer quadrant of the left breast and spot	0	1, 2, 3, 4, 5
10 Post-course test	3244705 and 3266705	Mass with spiculated margins and microcalcifications in the upper outer quadrant of the right breast	5	1, 4, 5

poor definition and sharpness of images when presented with the datashow device.

The capability of identifying the lesions, classification in accordance with BI-RADS, and coherence with recommended approach were evaluated.

RESULTS

The present study was aimed at improving the quality of mammographic diagnosis in three categories: lesion description, BI-RADS categories, and recommended approach. The average grades obtained between these three categories before and after the course did not show significant improvement in the learning, ranging from 5.6 to 5.8 in relation to lesion description, from 5.1 to 5.3 in relation to BI-RADS categories and from 5.5 to 5.9 in relation to recommended approach.

In relation to BI-RADS classification, categories assignment in the pre- and post-course tests for each mammographic study were compared with the following results:

1 – In post-course test a higher number of answers on the categories expected for those cases was observed. The only exception was observed for category 4, in which most students selected category 0, with no significant alterations, as the pre- and post-course testes were compared.

2 – Except for the correct category that should be chosen, the second most referred category in the cases was the same, as the pre- and post-course tests were compared (Table 2).

3 – As regards the correctness rate, there was only a small improvement in categories 2 and 4. In all the other categories, there was a reduction in the correctness rate as the pre- and post-course tests were compared.

4 – Category 3 was not utilized in the cases, considering that according to BI-

RADS the three alterations comprised in this category are: punctate, isodense and clustered microcalcifications (the course did not have appropriate viewing resources for microcalcifications evaluation), regular and well-defined masses, and focal asymmetry corresponding to confluent and nonpalpable breast tissue (these two alterations require additional mammographic evaluation, such as focal compression and magnification, which are not part of the baseline mediolateral oblique and cranio-caudal views, and were not among the images presented in the course)

Besides the evaluation of the classification according to the standardized system, the consistency regarding the recommended approach was also evaluated. In spite of the clarity of such system, primary inconsistencies were observed, for example:

- category 5 with six-month radiological follow-up;
- category 4 with six-month radiological follow-up;
- category 3 with recommendation for further evaluation with another imaging method;
- category 3 with 12-month radiological follow-up;
- category 3 with further pathological investigation;
- category 2 with recommendation for further evaluation with another imaging method;
- category 2 with six-month radiological follow-up;
- category 1 with recommendation for further evaluation with another imaging method;
- category 1 with recommendation for proceeding with pathological investigation;
- category 0 with recommendation for proceeding with pathological investigation.

In spite of such inconsistencies, it was observed that after the courses, many students only mentioned the BI-RADS category, without recommended approaches, or vice-versa.

Some frequent errors, such as description of axillary lymph nodes with usual radiological aspect as lymphadenopathies, diminished after the courses, particularly after the second one, considering that as the first group had already been analyzed, such

errors could be highlighted to the later groups.

As regards the analysis of breasts with silicone implants, there was a high rate of disagreement even in the post-course tests. Five attendees kept on selecting category 0, recommending further evaluation with ultrasonography, two attendees selected category 1 with a recommendation for 12-month radiological follow-up, one selected category 1 and another selected category 2 with recommendation for further evaluation with ultrasonography, and finally one attendee selected category 0, with recommendation for proceeding with pathological investigation.

DISCUSSION

The BI-RADS classification system, that is aimed at the standardization of radiological alterations and recommendations for approach to be adopted, has been well established and determined (Table 3).

In the evaluation of the capacity of identifying a mammographic lesion, its classification in compliance with BI-RADS and the consistency with recommended approach, the following points were observed:

1 – *Regular, well-defined mass (3167805)*: There was an improvement in the identification and characterization of the lesion, however there was no significant improvement in relation to BI-RADS classification and its consistency with the recommended approach. Some observed examples were: category 2 for mass; category 3 for mass with recommendation for further evaluation with other imaging method, or category 3 with 12-month radiological follow-up. It is important to note that one attendee did not correctly describe the alteration as a mass, but as an asymmetry; however he classified the lesion as category 0 with recommendation for further evaluation with another imaging method. In another case, a *regular mass with partially obscured limits (3401905)*, whose evaluation is a little more difficult than the first one, a greater difficulty in the identification was indeed observed; however, those who identified it continued to classify it into category 3, recommending six-month radiological follow-up. Other ex-

Table 2 Error in classification of BI-RADS categories in the presented cases.

Category	Second most selected category
0	3
1	2
2	1
4	0
5	4

Table 3 BI-RADS classification system.

BI-RADS category	Radiological finding	Recommended approach
0	Presence of a radiological finding requiring further evaluation with a complementary imaging method	Further evaluation with other imaging method is required
1	No radiological finding is observed. The study is normal	12-month radiological follow-up is suggested
2	Benign radiological findings	12-month radiological follow-up is suggested
3	Probably benign radiological findings (with < 2% chance of malignancy). This specific case includes clustered, punctate, isodense microcalcifications, regular, well-defined mass and nonpalpable focal asymmetry suggestive of confluent fibroglandular tissue	6-month radiological follow-up is suggested
4	Presence of radiological findings suspicious for malignancy – with a risk for malignancy ranging from 2% to 95%	Further diagnostic investigation is required. Bi-opsy is recommended
5	Highly suspicious radiological findings, with a malignancy risk greater than 95%	Further diagnostic investigation is required. Bi-opsy is recommended
6	Proven malignant lesion	—

amples were category 4 with recommended histopathological investigation or category 0 with recommendation for further investigation with other imaging method.

2 – *Highly suspicious high-density spiculated mass (3244705 and 3266705)*: The mass was correctly described by the attendees, however some of them did not mention the associated findings, but most classified the lesion as category 5 and some as category 4, with recommendation for proceeding with histopathological investigation. This is a relevant fact, as in such case what really matters is that a highly suspicious mass be histologically evaluated.

3 – *Architectural distortion (3177605 and 32688805)*: In the first case the architectural distortion was more subtle, with slight improvement when the pre-course and post-course tests were compared, including the description of other lesions that were not present and that guided the classification in compliance with BI-RADS, and the recommended approach, which in most cases was not consistent; in the second case, in which the distortion was more noticeable, with associated findings increasing the suspiciousness of the lesion, the descriptions remained practically unchanged in the pre- and post-course tests (not all attendees described the architectural distortion, some described it as a mass or as a distortion associated with a mass), as well as in the BI-RADS categorization; however, the identification of the lesion malignancy characteristics was more evident, as there was a significant improvement with respect to recommendation for

proceeding with histopathological investigation.

4 – *Benign calcifications – with lucent center, round calcifications (3252005) and vascular calcifications (3283405)*: In the first case, no significant improvement was observed in spite of the fact that the calcifications were in general mostly identified and described correctly; inexistent alterations were described, which conditioned the BI-RADS classification and approach recommendations that were not always consistent. On the other hand, in the case of vascular calcifications, no improvement was observed in the characterization, with many attendees persisting in the same description, which included inexistent alterations, consequently with wrong categories and recommended approach.

5 – *Operated breasts (silicone implants-3256005 and reduction mammoplasty-3407905)*: All the attendees identified the implant, with the exception of one that only described it in the post-course test. No significant improvement was observed with regards to BI-RADS classification and recommended approach. Two facts should be highlighted: some attendees identified inexistent lesions which compromised the BI-RADS classification and recommended approach, and the fact that some believed that breasts with silicone implants required other imaging methods for a thorough study. With respect to the case with reduction mammoplasty, only four attendees identified the case as operated breasts, and only these attendees continued describing the reduction mammoplasty in the post-

course test, with no comparative improvement in the pre- and post-course tests; thus, for this particular case, the BI-RADS classification and recommendation for approach were based on alterations that were not present. It should also be mentioned that in many cases the BI-RADS classification was not consistent with the recommended approaches.

6 – *Normal studies presenting processing errors – roll marks (3229205)*: There was no significant improvement with the course. Only the same attendees that identified the study as normal in the pre-course test, continued to do so in the post-course test; however, in most groups inexistent lesions were identified, which obviously misdirected the BI-RADS classification, which moreover, was many times not consistent with the recommended approach. Roll marks were poorly observed, and only in the pre-course test.

Even after some changes in the course structure, with direct approaches and use of repetitive methods, no significant improvement in level of learning was observed.

It was observed that many times the attendees visualized the radiological alteration, but could not describe it correctly; frequently, confusion between focal asymmetry and mass was observed, as well as the description of architectural distortion as a mass with malignancy characteristics. It should be pointed out that the most relevant aspect is the identification of the radiological alteration and the approach to be adopted by the assisting physician for the patient.

Many times, in cases classified as category 0, it is possible to define that a mass is probably benign after all the additional views required for the study show a nonpalpable regular mass.

Categories 1 and 2 correspond to mammographic studies with recommendation for a 12-month follow-up, considering that such studies with results negative for malignancy with no harm for the patient.

Categories 5 and 4 indicate respectively highly suspicious and suspicious mammographic findings; in such cases, recommendation of histopathological investigation of the described finding is mandatory. Therefore, when category 5 is confused with category 4, the patient will not suffer any harm, since in both categories the recommended approach will be proceeding with histopathological diagnosis. However, when a radiologically suspicious lesion is classified as category 0, the correct diagnosis will be delayed, as a further imaging study, ultrasonography for example, would not be necessary to confirm a suspicious finding that could have been referred for biopsy after the initial mammographic study.

Berg et al.⁽⁸⁾ have developed a study to determine whether training on BI-RADS improves observer performance and interobserver agreement among experienced breast radiologists, taking into consideration the analysis of mammographic characteristics and the final evaluation. Fifty-four lesions were selected: 28 nodules (9 malignant) and 26 microcalcifications (10 malignant). The observer performance was evaluated before and after the BI-RADS training course and there was improvement in agreement regarding the description of mass margins. A similar improvement was observed in the description of morphologic characteristics of microcalcifications. No improvement was observed in the description of calcifications distribution. The final evaluations were more consistent after the training. Recom-

mendations for biopsy increased from 73% to 88%, with a subtle increase in the number of recommendations for biopsy in benign lesions (from 43% to 51%).

Sickles et al.⁽⁹⁾ have observed that radiologists specialized in breast imaging undergo courses with a six-times higher frequency than generalist radiologists, and interpret a ten times higher number of mammograms per year, being capable of diagnosing a greater number of cases of early stage breast cancer, recommending more biopsies and presenting lower rates of patient recall than generalist radiologists.

CONCLUSIONS

1 – Breast radiologists in the public health system of the Rio de Janeiro State demonstrate insufficient knowledge on early imaging diagnosis of breast cancer.

2 – The course for training medical practitioners in the early detection of breast cancer, with emphasis on theoretical classes, did not produce a significant improvement in the quality of mammography interpretation, with persistent mistakes in morphological description of critical breast lesions and in their classification by the BI-RADS, as well as poor consistency between the BI-RADS classification and the recommended approach.

Recommendations

1 – The training courses must emphasize the practice by means of mammographic cases in which the repeated demonstration of critical breast lesions, their BI-RADS classification and recommended approaches are thoroughly covered and trained.

2 – Diagnosis errors can be minimized when the characteristics of typically benign or malignant lesions are well defined, reducing the uncertainty in mammographic diagnosis.

3 – The knowledge acquired in theoretical courses is important. However, one

must bear in mind that advanced technologies (such as magnetic resonance imaging in the evaluation of breast diseases) are not widely utilized in Brazil. However, programs of theoretical courses for breast cancer diagnosis emphasize such novel diagnosing methods.

4 – It is necessary to promote the gathering of professionals involved in the teaching of mammographic diagnosis in order to reach a consensus on a course curriculum and the most appropriate teaching method.

REFERENCES

1. Elmore JG, Carney PA. Does practice make perfect when interpreting mammography? *J Natl Cancer Inst.* 2002;94:321–3.
2. Elmore JG, Nakano CY, Koepsell TD, et al. International variation in screening mammography interpretations in community-based programs. *J Natl Cancer Inst.* 2003;95:1384–93.
3. Eiras AL, Koch HA, Peixoto JE. Parâmetros envolvidos na qualidade da imagem mamográfica – revisão dos fundamentos teóricos. *Rev Imagem.* 2000;22:143–8.
4. Koch HA. Diagnóstico por imagem mamária no Brasil. Rio de Janeiro: Academia Nacional de Medicina; 2004.
5. Linver MN, Paster SB, Rosenberg RD, et al. Improvement in mammography interpretation skills in a community radiology practice after dedicated teaching courses: 2-year medical audit of 38,633 cases. *Radiology.* 1992;184:39–43.
6. Colégio Brasileiro de Radiologia, Sociedade Brasileira de Mastologia, Federação das Sociedades Brasileiras de Ginecologia e Obstetrícia. I Reunião de Consenso da Radiologia – Padronização dos laudos mamográficos. 28ª Jornada Paulista de Radiologia, São Paulo, abril de 1998.
7. Luna MM. Avaliação dos laudos mamográficos: padronização de recomendação de conduta para um programa de detecção precoce do câncer de mama por meio da mamografia [tese de doutorado]. Rio de Janeiro: Universidade Federal do Rio de Janeiro; 2001.
8. 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.
9. Sickles EA, Wolverson DE, Dee KE. Performance parameters for screening and diagnostic mammography: specialist and general radiologists. *Radiology.* 2002;224:861–9.