Accuracy of the Mammographic Report Category according to BIRADS™

Siriporn Hirunpat, MD*, Wiwatana Tanomkiat, MD*,
Rungrat Khojarern, MD*, Nualta Arpakupakul, MSc, MA**

** Department of Radiology, Faculty of Medicine, Prince of Songklanagarind Hospital University*
*** Epidemiology Unit, Prince of Songklanagarind Hospital University*

Objective: *To determine the accuracy and predictive value of the mammographic report according to the BIRADS categories in Songklanagarind University Hospital.*
Material and Method: *Mammograms of 1000 women who came to Songklanagarind University Hospital from June 1998 to September 1999 were reported and placed in category 1 for negative results, category 2 for benign lesions, category 3 for probably benign lesions, category 4 for suspicious lesions and category 5 for highly suspicious lesions. Accuracy was determined by either histology or by unchanged follow-up mammography within 24 months.*
Results: *The total accuracy was 97.8%, sensitivity 62.5% and specificity 98.1%.*

Keywords: *Mammographic, Breast cancer, BIRADS*

J Med Assoc Thai 2005; 88(1): 62-5
Full text. e-Journal: *http://www.medassocthai.org/journal*

Mammography is a well known and effective tool for detecting early breast cancer. Prince of Songkla University Hospital, which is the biggest university hospital in Southern Thailand serving approximately 10 million people in 14 provinces of the Southern part of Thailand, has been performing mammography for women since 1994. The initial report system was free style according to the opinion of each radiologist.

The American College of Radiology established the Third Edition Guidelines for the Breast Imaging Reporting and Data System (BI-RADS™) in 1998 in order to standardize the terminology in the mammographic reports on the basis of the level of suspicion.

This report system has been used in Songklanagarind University Hospital since 1999 and is accepted to be a very useful and practical way of communication between radiologists and clinicians. However, the accuracy and the predictive values of the mammographic reports according to the BI-RADS categories in Songklanagarind University Hospital have not yet been studied. The purpose of the present study was to determine the accuracy of the mammography reports and calculate the predictive values of each BIRADS category in the authors' mammographic reports in Prince of Songkla University Hospital.

Correspondence to : Hirunpat S, Department of Radiology, Faculty of Medicine, Prince of Songklanagarind University, Songkla 90110, Thailand.

Material and Method

Data was retrospectively collected from the medical records, pathologic records and mammograms of 1000 women who came to Songklanagarind University Hospital for either screening or diagnostic mammography to exclude possible breast cancer from June 1998 to September 1999. At least two standard views were taken on each breast (one craniocaudal view and one mediolateral oblique view) by a dedicated mammography machine (Toshiba MGU-10 A). The first-visit films and reports of every patient were retrospectively re-classified into one category according to BIRADS by an experienced general radiologist in Prince of Songkla University Hospital who was blinded from both clinical and pathologic findings. Negative studies were placed into category 1 and mammographically detected lesions were placed into one of the four BI-RADS assessment categories: category 2 for benign lesions, category 3 for probably benign lesions, category 4 for suspicious lesions and category 5 for highly suspicious lesions for malignancy.

The accuracy of the radiologist's report was determined by either histology (biopsy, excision or lumpectomy) or by unchanged mammography. To be diagnosed as having no cancer, all patients who had no proved histology must have an unchanged follow-up mammography within 24 months. Any mammographic change must be histologically proven and the period time between the abnormal mammogram and the histological specimen had to be within 60 days. Patients who had no histological specimens to prove either the benign or malignant nature of the lesion, or were lost to follow-up mammography before 24 months were excluded from the study. The positive predictive value, negative predictive value, sensitivity, specificity and accuracy were calculated.

Results

There were 870 patients, 365 cases of screening mammography and 505 cases of diagnostic mammography, who fulfilled the inclusion criteria. The patients' age ranged from 20-79 years, with an average age of 46 years.

The mammographic findings were categorized according to BIRADS as follows;

BIRADS 1	476	54.71%
BIRADS 2	331	38.04%
BIRADS 3	42	4.82%
BIRADS 4	16	1.83%
BIRADS 5	5	0.57%

Histological findings in 18 of 476 patients (3.8%) who were categorized as BI-RADS 1 revealed fibrocystic disease (7), fibrosis (5), fibroadenoma (3), chronic inflammation (1), unremarkable specimen (1) and infiltrative ductal carcinoma (1).

The other 458 patients showed an unchanged follow-up mammogram in a period of 24 months.

Histologic findings in 28 of 331 patients (8.5%) who were categorized as BIRADS 2 revealed fibrocystic disease (13), fibroadenoma (13), abscess (1) and infiltrative ductal CA (1). The remaining 303 patients showed an unchanged follow-up mammogram in a period of 24 months.

Histological findings in 9 of 42 patients (21.4) who were categorized as BIRADS 3 revealed fibrocystic disease (5), fibroadenoma(1), chronic inflammation (1), abscess (1), and infiltrative ductal CA (1). The remaining 33 patients showed an unchanged follow up mammogram within 24 months.

Histologic findings in 10 of 16 patients (62.5) who were categorized as BIRADS 4 revealed fibrocystic disease (4), fibroadenoma (2), fibrosis (2) abscess (1) and infiltrative ductal CA (1). The remaining 6 patients showed an unchanged follow up mammogram within 24 months.

All 5 patients who were categorized as BIRADS 5 had had surgery and were histologically proven to be infiltrative ductal carcinoma in 4 cases, and the other, a benign lesion, was an abscess in a patient who had surgical drainage 2 weeks before the mammograhy but still had a palpable mass lesion.

BIRADS 1, 2 and 3 were regarded as negative studies and BIRADS 4 and 5 were regarded as positive studies for malignancy[1]. Therefore, there were 849 negative studies and 21 positive studies (Table 1).

True negative (TN) in BIRADS categories 1,2, and 3 was regarded as no known diagnosis of cancer within 2 years of mammographic examination or benign biopsy findings within 60 days after the abnormal mammographic examination (BIRADS categories 2 and 3).

TN was 475 + 330 + 41 = 846

True positive (TP) was regarded as cancer diagnosed within 60 days of the mammographic examination categories 4 and 5.

TP was 1 + 4 = 5

False negative (FN) was regarded as diagnosis of cancer within 24 months of mammographic examination with normal or probably benign findings (BIRADS categories 1, 2 and 3).

FN was 1 + 1 + 1 = 3 cases

False positive (FP) was no proven cancer diagnosis within 2 years of the positive mammographic examination study (BIRADS categories 4 and 5) or benign biopsy findings within 60 days after the mammographic examination with BIRADS categories 4 and 5.

FP was 15 + 1 = 16 cases

Negative predictive value = TN/number of negative mammographic examinations
= 846/849
= 99.64%

Table 1. Negative and Positive predictive values separately calculated for each BIRADS category

BIRADS	NPV	PPV
1	99.8% (475/476)	
2	99.7% (330/331)	
3	97.6% (41/42)	
4		6.2% (1/16)
5		80.0% (4/5)

Positive predictive value = TP/number of positive mammographic examinations
= 5/21
= 23.8%

Sensitivity = The probability of detecting a cancer when a cancer existed or the number of cancers diagnosed with mammographic examination, in a population within 2 years of their imaging examinations.

Sensitivity = TP/TP + FN
= 5/5 + 3
= 62.5%

Specificity = The number of mammographically normal cases in a population divided by all normal cases in the population.

Specificity = TN/TN + FP
= 846/846 + 16
= 98.1%

Accuracy = TN+TP/All patients
= (846 + 5)/870
= 97.8%

Discussion

Mammography has been widely used as the most important screening and diagnostic tool for breast cancer. BIRADS is useful to standardize mammographic reports and has been accepted as an effective and practical reporting system both for radiologists and clinicians. The final assessement category gets rid of the confusion about the further management of the patients. Mammographic reports with BIRADS categories 1, 2 and 3 were classified as negative. Mammographic reports with BIRADS categories 4 and 5 were classified as positive.

In spite of the lack of a mammographic expert in Songklanagarind University Hospital, the general radiologist did quite well for imaging interpretation with a sensitivity of 62.5%, specificity of 98.1% and overall accuracy of 97.8%. The specificity was similar to previous reports(93.4%, 80.7%)[2,3]. However, the sensitivity was much less (89.83%, 84, 86.6%)[2-4]. The rather low sensitivity might be partly from the small number of positive cases in our study population[8] and partly from the interpretation by the general radiologist who has interpretated only about 10-20 mammography cases per week for 6 years. It has been reported that the more experienced a radiologist, the more the sensitivity in diagnosing breast cancer. Reader volume is also an important determinant of mammogram sensitivity and specificity. High sensitivity and high specificity can be achieved in high-volume centers[5,6].

The high overall negative predictive value (99.6%) and low positive predictive value (23.8%) reflected the low prevalence of the breast cancer in the present study population. There were only 8 cases in 870 patients; 2 cancers from 365 screening cases and the other 6 cases from 505 diagnostic cases. The low prevalence of breast cancer was similar to other parts of Thailand, which was 665 per 100,000 population in 1996[7]. However, this overall positive predictive value was within the accepted range for the field of 15-30% [8,9].

Another reason that might result in a low PPV could be that the authors included both BIRADS 4 and 5 in the same positive mammographic examination group. A separate calculation of each category might be more reasonable and improve the quality of risk assessment[10]. With a seperate calculation, the PPV for BIRADS 4 was 6.25% (1/16) and PPV for BIRADS 5 was 80% (4/5). The separate calculation of the negative predictive value for BIRADS 3 was 97.63% (41/42).

In contrast to other reports[10,11] in which PPV were calculated from the surgical base population and most cases were BIRADS 4 and 5, the present study determined the predictive values in both a non-surgical and surgical population most of which were BIRADS 1, 2 or 3. Negative predictive value was calculated for BIRADS 1, 2 and 3 instead of PPV. A high negative predictive value was achieved in BIRADS 1, 2 and 3 regardless of whether the patients had symptoms or not. With this way of calculation, the positive predictive value can be calculated for only BIRADS 4 and 5 but not for BIRADS 0, 1, 2, or 3 as in other reports[10,11].

Similar to other reports[11], the most common histologic diagnoses of benign lesions biopsied in our institute were fibroadenoma and fibrocystic changes, and the histologic diagnoses of malignant lesions in the present study group were infiltrating ductal carcinoma. However, not a single case of carcinoma in situ was found in the present study.

All 4 breast cancer patients within BIRADS categories 1, 2, 3 and 4 came to Songklanagarind University Hospital due to palpable mass lesions. This supports the truth that, not all palpable lesions will be detected by mammography. In spite of the likelihood of a negative or benign mammographic appearance, there is still a small possibility of malignancy. The other 4 cancer patients within BIRADS category 5 came to Songklanagarind University Hospital with palpable masses in 2 cases and for screening for cancer in the other 2 cases, which reflects the usefulness of mammography and could detect 2 cases out of 365

screening patients in the present study. However, the cost- effectiveness of screening mammography in Thailand should be further determined. Therefore, the combination of physical examination and mammography was again confirmed to be the best way of early breast cancer detection.

References

1. Breast Imaging Reporting and Data System BIRADS 3rd ed. American College of Radiology 1998.
2. Charnvises S. The accuracy of mammography: a comparison with fine needle aspiration cytology. J Med Assoc Thai 1996; 79: 27-30.
3. Manoliu RA, Ooms GH. The accuracy of mammography: an analysis of 655 histologically verified cases. Radiol Clin (Basel) 1977; 46: 422-9.
4. Azzarelli A, Guzzon A, Pilotti S, Quagliuolo V, Bono A, Di Pietro S. Accuracy of breast cancer diagnosis by physical, radiologic and cytologic combined examinations. Tumori 1983; 69: 137.
5. Elmore JG, Wells CK, Howard DH. Does diagnostic accuracy in mammography depend on radiologists' experience? J Womens Health 1998; 7: 443-9.
6. Esserman L, Cowley H, Eberle C, et al. Improving the accuracy of mammography: volume and outcome relationships. J Natl Cancer Inst 2002; 94: 369-75.
7. World Health Organization. Women's health and development country profile: Thailand. New Delhi: WHO Regional Office for South-East Asia, 1998.
8. Kopans DB. The positive predictive value of mammography. AJR 1992; 158: 521-6.
9. Adler DD. Helvie MA. Mammographic biopsy recommendations.Curr Op Radiol 1992; 4: 123-9.
10. Orel SG, Kay N, Reynold C, Sullivan DC. BI-RADS categorization as a predictor of malignancy. AJR 1999; 211: 845-50.
11. 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.

ความถูกต้องแม่นยำของการรายงานผลแบบไบแรดส์ของการตรวจเต้านมด้วยแมมโมแกรม

สิริพร หิรัญแพทย์, วิวัฒนา ถนอมเกียรติ, รุ่งรัตน์ ขอเจริญ, นวลตา อาภาคัพพะกุล

วัตถุประสงค์*: เพื่อหาค่าความถูกต้องแม่นยำของการรายงานผลแบบไบแรดส์ของการตรวจเต้านมเพื่อค้นหามะเร็งด้วยแมมโมแกรมในโรงพยาบาลสงขลานครินทร์*

วิธีการ*: การศึกษาในหญิงที่ได้รับการตรวจแมมโมแกรมในโรงพยาบาลสงขลานครินทร์จำนวน 1000 ราย ระหว่างมิถุนายน 2541 - กันยายน 2542 ผลการตรวจที่ไม่พบความผิดปกติจัดเป็นไบแรดส์ชนิดที่ 1 ผลการตรวจที่พบความผิดปกติที่ไม่ใช่มะเร็งจัดเป็นไบแรดส์ชนิดที่ 2 ผลการตรวจที่พบความผิดปกติที่น่าจะไม่ใช่มะเร็งจัดเป็นไบแรดส์ชนิดที่ 3 ผลการตรวจที่พบความผิดปกติที่สงสัยว่าอาจเป็นมะเร็งจัดเป็นไบแรดส์ชนิดที่ 4 ผลการตรวจที่พบความผิดปกติที่สงสัยเป็นอย่างยิ่งว่าอาจเป็นมะเร็งจัดเป็นไบแรดส์ชนิดที่ 5 หาความถูกต้องแม่นยำโดยผลการตรวจชิ้นเนื้อ ส่วนรายที่ไม่มีผลชิ้นเนื้อ จะต้องได้รับการตรวจแมมโมแกรมต่อไปจนครบ 24 เดือน*

ผล*: ความถูกต้องแม่นยำของการรายงานผลร้อยละ 97.81 ความไวร้อยละ 62.5 ความจำเพาะร้อยละ 98.14*
