BI-RADS OUTCOME ASSESSMENT OF MAMMOGRAPHY SCREENING; MEDICAL AUDIT OF A BREAST **IMAGING CENTER**









¹Altunizade Acıbadem Hastanesi, Meme Kliniği, İstanbul, Türkiye

²Acıbadem Mehmet Ali Aydınlar Üniversitesi, Tıp Fakültesi, İstanbul, Türkiye

³Sancaktepe Şehit Prof. Dr. İlhan Varank Eğitim ve Arastırma Hastanesi, Radyoloji Bölümü, İstanbul, Türkiye

⁴Acıbadem Mehmet Ali Aydınlar Üniversitesi, Tıp Fakültesi Radyoloji Anabilim Dalı, İstanbul, Türkiye

Nilgün GÜLDOĞAN, MD Damla GÜRLEYİK Fikret KÜÇÜKKAYA, MD Erkin ARIBAL, MD

İletişim: Damla Gürleyik Acıbadem Mehmet Ali Aydınlar Üniversitesi, Tıp Fakültesi, İstanbul, Türkiye Tel: +902165004444

E-Posta: damlagurleyik@yahoo.com

Gönderilme Tarihi: 29 Mart 2019 **Revizyon Tarihi**: 08 Ekim 2019 Kabul Tarihi : 13 Ekim 2019

MAMOGRAFI TARAMA SONUÇLARININ BI-RADS'A GÖRE DEĞERLENDIRİLMESİ; BİR MEME GÖRÜNTÜLEME MERKEZİNİN TIBBİ DENETİMİ

Amaç: Kliniğimizde 2017 ve 2018 yıllarında yapılan tarama mamografilerinin sonuçlarının American College of Radiology (ACR) önerileri doğrultusunda karşılaştırmalı olarak değerlendirmeyi amaçladık.

Yöntem: 14 Mart 2017 ve 31 Aralık 2018 tarihleri arasında yapılan tarama mamografi sonuçlarını kaydedip analiz ettik. Bu analizleri kullanarak tarama mamografi uygulamalarımızı denetledik. Sonuçlar ACR BI-RADS önerileri ve performans kriterlerine göre değerlendirildi.

Bulgular: Meme kanseri öyküsü bulunmayan kadınlara yapılan tarama mamografisi sayısı toplam 2413'tü. Pozitif Öngörü Değeri (PÖD)1, PÖD2 ve PÖD3 hesaplamalarımız sırasıyla %5,6, %30,8 ve %52,9 idi. Geri çağırma oranı ise %11'di. Kanser saptama oranı, minimal kanser oranı, erken evre kanser oranı ve lenf nodu negatifliği çıktıları sırasıyla %0,74, %50, %77 ve %64 olarak saptandı.

Sonuc: Tarama mamografilerinin değerlendirme çıktılarının ACR performans kriterleri ile uyumlu olduğu görüldü. Erken evre meme kanseri saptama oranı istenildiği gibi yüksek olmasına rağmen aksiller lenf nodu negatiflik oranının istenilen %75'lik oranın altında kaldığı görüldü. Medikal denetleme, meme görüntüleme biriminin tarama ve tanı çalışmalarının etkinliğini gösterme ve birimin uygulamalarının iyileşmesi için etkili bir geri bildirim yöntemi olması açısından önemlidir.

Anahtar sözcükler: Medikal denetleme, mamografi, ultrason, otomatik meme ultrasonu, tarama

ABSTRACT

Aim: To audit our mammography screening practice according to the American College of Radiology (ACR) recommendations in the years 2017 and 2018.

Method: We documented and analyzed our breast imaging data collected between March 14th, 2017 and December 31st, 2018 and performed a medical audit related to our mammography screening practice. The results were evaluated according to ACR BI-RADS recommendations and performance benchmarks.

Results: The total number of mammographic screening patients who had no personal history of breast cancer was 2413 in 2017 and 2018. Our PPV1, PPV2 and PPV3 values were 5.6%, 30.8%, 52.9% accordingly. The recall rate was 11%. Outcomes for cancer detection rate, minimal cancer rate, early-stage cancer rate and lymph node negativity were 0.74%, 50%, 77% and 64% respectively.

Conclusion: The mammographic screening outcomes are in accordance with ACR performance benchmarks. Despite a desired high early cancer detection rate, axillary lymph node negativity was found to be below the desired range of 75%. Medical auditing is important for a breast imaging unit to monitor the effectiveness of screening and diagnostic procedures and presents as an effective feedback tool to improve the practice.

Keywords: Medical audit, mammography, ultrasound, automated breast ultrasound, screening

n audit is the interpretation of the data collected in a certain period of time. Audit has become one of the main growth areas in the medical literature since the 1970s (1). It is a tool for measuring the quality of medical practice in order to improve standards in clinical care and encourage the efficient utilization of resources. For mammogram; the purpose of a medical audit is to provide feedback to the facilities and doctors on their performances relative to established benchmarks and to improve the overall quality of the breast cancer screening practice. It gives an assessment of one's ability to detect occult breast cancer, which is the main indicator of mammography screening performance. Auditing evaluates the percentage of cancers, ratio of curable stage cancers among them and additional diagnostic procedures including the biopsies recommended (2,3)

Auditing does not only give an idea about the performance of the diagnostic practice but also points out any need for corrective action. An outcome below the expected range indicates a need for additional education or changing in the practice setting. If the results are within the expected ranges, the confidence of the interpretations and compliance of the referring physicians and patients can be confirmed.

Poor quality mammography may result in missed cancers, false-positive examinations, increased costs and anxiety. Quality assurance standards have been promoted after the recognition of the importance of quality control of screening mammography. The Mammography Quality Standards Act (MQSA) in the USA required all mammography facilities to become accredited and certified in 1994. This has led to a significant improvement in mammogram quality in the USA over the last 20 years. However, in Turkey, there is no mandatory accreditation program.

Recent studies in Turkey showed that quality control or precautions are not taken seriously (4). The purpose of this breast imaging audit is to provide feedback on the performance of the screening program of the patients without a personal history of breast cancer and compare these results with established benchmarks (5.6)

The breast imaging unit was founded on the 14th of March 2017 when we started accepting screening patients. Since the opening of our breast imaging department, we have been collecting and documenting patients' data for auditing purposes. In this study, we aim to present a detailed audit data collected in our breast imaging department in

2017 and 2018 and evaluate it according to the American College of Radiology (ACR) and Breast Imaging Reporting and Data System (BI-RADS) recommendations and performance benchmarks and to put forth the efficacy of auditing in initializing a breast imaging center.

MATERIAL AND METHOD

Screening mammography is performed in asymptomatic women. The onset age for screening is 40 years according to the recommendation of the National Screening Programme. The inclusion criteria for auditing were: i: women without any symptoms or signs of breast disease, ii: women equal or over age 40, iii: women who had mammography examination for screening. The exclusion criteria were: i: women who had a history of self-breast cancer, ii: women who had symptoms for breast abnormality, iii: women who did not have a mammogram for screening but preferred US screening, iv: women younger than 40 years, v: women whose findings were failed to be documented with our manual documentation. This study is approved by our institutional review board.

The numbers of mammograms performed were 978 and 2124 in 2017 and 2108 respectively. Of these examinations that have met the inclusion criteria, 724 and 1689 were screening mammograms for years 2017 and 2018 respectively with a sum of 2413. The number of women that were not documented was 359. The mean age of women included in this study was 50.47. The number of women between the ages of 40 and 49 years was 920 (44.38%) with a mean age of 44.08 and the number of women above 50 years of age was 1153 (55.62 %) with a mean age of 58.47.

Every patient referred to our breast imaging department was asked to complete a questionnaire in which medical history, personal and family history of breast cancer, history of biopsies and surgical procedures of the breast as well as their demographic data were documented. All patients' 5 years and lifetime risk were evaluated according to the Gail model and in cases with a familial cancer history, the Claus model was used. The lifetime risk was assessed as low when it was lower than 15%, moderate between 15-20% and high when equal to or above 20%. The 5-year risk was assessed as low when below 1.7% and high when equal to or above 1.7%.

All patients were examined with two-view mammograms (Senographe PristinaTM, GE Healthcare). The craniocaudal images were taken with 2D imaging and mediolateral

oblique (MLO) images with tomosynthesis. Synthetic 2D MLO views and tomosynthesis slices were used for the evaluation of MLO acquisitions. The women who had free combined mammography and US screening option according to their insurance policy had both screening examinations. The US screening was done either with a handheld US (HHUS) (Logic S8, GE Healthcare) or an automated breast ultrasound system (InveniaTM ABUS, GE Healthcare). A linear 9-16 MHz linear transducer probe was used for HHUS and each breast was scanned with both radial and continuous transverse techniques. ABUS examination was performed with a 15 cm wide 6-15 MHz reverse curved standard probe in at least 3 projections (AP, Lateral and Medial). Additional necessary projections were added in large breasts in order to cover the whole breast area. All cases were evaluated and reported by one of the two experienced breast radiologists (one experienced in breast radiology for 6 years and the other one for 20 years) according to BI-RADS recommendations. Double reading was started after November 2018. The breast density of each case was evaluated visually according to BI-RADS definitions and recorded for each patient (2). Most of the women in our clinic had mammography and US screening as a combo screening examination according to their insurance screening policy. In this case, all mammography examinations were categorised independently from the US screening findings. A normal mammogram was evaluated as BI-RADS 1 even in the existence of a positive finding on the US screening. On the other hand, BI-RADS 0 was only used in the need of an additional diagnostic examination due to an abnormal finding on the mammogram.

All data related to the BI-RADS category and breast density are recorded manually. Tumor size and type, lymph node status and cancer stages were added after the pathology results of the breast cancer cases were obtained. The raw data included audit period dates, number of mammographic and ultrasound examinations, number of recalls, number of recommended biopsies, biopsy results and tumor staging. We calculated the below outcomes derived from the raw data (2,6,7) according to BI-RADS recommendations:

Positive predictive values (PPV1, PPV2, PPV3):

PPV1: Indicates abnormal exam based on positive cases at screening examination, which includes a recommendation for anything other than routine screening (BI-RADS® categories 0,3,4,5)

PPV2: Number of biopsies recommended (BI-RADS® categories 4, 5)

PPV3: Based on results of biopsies that actually performed, also known as biopsy yield of malignancy or positive biopsy rate.

Percentage of cancers that are "minimal": minimal cancer is defined as invasive cancer ≤ 1 cm, or ductal carcinoma in situ [DCIS] of any size (8,9).

Percentage of cancers that are stage 0 and 1. Staging of cancers is based on the manual of the American Joint Committee on Cancer (AJCC) (8).

Percentage of invasive cancers that are node-negative. The percentage of axillary lymph node-negative cases with invasive cancer in ratio to all invasive cancer patients.

Cancer detection rate is the number of cancers detected per 1000 women examined.

Recall rate is the percentage of cases that are recommended for any further imaging evaluation. It shows the cases with BI-RADS 0 assessment category. We compared our results with the standard recommended ranges defined in ACR BI-RADS ATLAS 5th edition.

RESULTS

The total number of patients who applied to the breast radiology department was 5867 (2188 in 2017 and 3679 in 2018). In 2017, 359 patients out of 3679 women, (16.4%) were not documented, due to the loss of documents during the manual data entering process. All the patients were documented in 2018.

The number of screening mammography was 724 and 1689 (sum=2413) in 2017 and 2018 respectively. Visual breast density assessment of screening mammograms was as follows: 9.60% category A, 31.2% category B, 38.6% category C, 20.6% category D. The rate of women who had dense mammograms (C and D type) was 59.2%. The density distribution according to age groups is given in Table 1.

The lifetime and 5-year risk findings of women according to age groups are given in Table 2. The screening modalities like mammogram only, mammogram and HHUS or mammogram and ABUS with correlating BI-RADS category findings are given in Table 3.

Table 1: Breast density distribution according to different age groups.					
Breast density (ACR)	A	В	С	D	
Age					
40-49	4.3%	22.9%	41.5%	31.3%	
50-59	10.8%	40.3%	39.6%	9.3%	
60<	24%	43.1%	28.4%	4.4%	
Total	9.6%	31.2%	38.6%	20.6%	

Table 2: Lifetime and 5-year risk assessments of patients according to different age groups					
	Lifetime Risk assessment			5 Year Breast Cancer Risk	
Age	<15 Low	15≤ <20 Moderate	≥20 High	<1.7 Low	≥1.7 High
40-49	88.7%	9.1%	2.2%	94.5%	5.5%
50-59	92.7%	6.4%	0.8%	84.5%	15.5%
60<	98.2%	1%	0.8%	50.3%	49.7%
Total	91.4%	7%	1.6%	84.2%	15.8%

Table 3. BI-RADS ass years 2017 and 201		nography and ultrasound	in the screening of pat	ients with no personal	history of breast cancer for	
	Combined screet	Combined screening MG+US (2017= 90.5%) (2018= 84%)				
2017						
BI-RADS	MMG	HHUS	ABUS	All US	MMG	
0	76	12	9	21	10	
1-2	498	345	229	574	55	
3	70	15	14	29	3	
4	10	10	17	27	1	
5	1	2	2	4	-	
Sum	655	384 (58.6%)	271 (41.4%)	655 (100%)	69	
2018						
0	150	19	27	46	41	
1-2	1104	387	845	787	200	
3	130	68	17	85	22	
4	31	25	24	49	7	
5	3	2	4	6	1	
Sum	1418	501 (35.3%)	917 (64.7%)	1418 (100%)	271	
2017+2018	2073	885	1188	2073 (85.9%)	340 (14.1%)	
Cancers Detected		15			3	
MMG: mammograph	າy, HHUS: handheld ເ	ultrasound, ABUS: automat	ed breast ultrasound			

Negative or benign assessment (BI-RADS category 1 and 2) was found in 77% of examinations performed in 2017 and 2018. The recall rate (BI-RADS 0 assessment) was 11%. Three cancers were detected with mammography screening among 724 cases in 2017. Two of them were detected with both modalities (US and mammogram). One cancer was detected with mammography screening per se. On the other hand, 17 cancers were detected in 2018. The mammography and US detected 15 of 17 cancers separately. Two cancers were detected only by mammography while another 2 only by ultrasound. One of the two cancers was detected only with US and had BI-RADS category 1 and the other one category 2 was detected with mammogram and had dense mammograms (both were graded as C type fibroglandular tissue composition). Cancers detected by screening mammography are shown in Table 4.

Clinical outcomes, performance measurements and cancer outcomes are reflected in Table 5. A biopsy was recommended for 52 cases and 18 of these (34.6%) were lost in follow-up. The number of biopsies performed was 34 and cancer was detected in 18 (52.9%) of these biopsies. Among these 18 detected cancers 4 were DCIS, 9 were minimal cancers (%50), 9 of the invasive cancers were axilla negative (%64) and 77% of cancers were stage 0 or 1 with a negative axilla.

Table 4. Cancers detected in the screening group in 2017 and 2018				
	# CASES			
DCIS	4			
Invasive cancers	14			
Less than 1cm	5			
1-2 cm	5			
Greater than 2cm	3			
Metastatic	1			
Axilla (-) invasive cancers	9			
Total	18			

Table 5: Clinical and performance outcomes of mammographic screening patients					
	2017 and 2018	ACR	Tunçbilek	Turk	Kayhan
Number of Mammographies	2413		6858	3048	3758
Biopsy recommended	52		122	105	
Biopsy performed	34		103	70	
Cancers detected	18		42	21	18
PPV1	5.6%	5-10%	4.9%	9.9%	2.3%
PPV2	30.8%	20-40%	33%	20%	26.9%
PPV3	52.9%	25-45%	39%	25%	32.7%
Cancer detection rate	0.74%	≥0.25%	0.61%	0.69%	0.45%
Recall rate	11%	5-12%	10.9%	3.5%	18.4%
% of node (-) invasive cancers	64%	75%	76%	66.7%	83.4%
% of minimal cancers	50%	>30%	38%	14.3%	50%
% of stage 0-1 cancers	77%	>50%	50%	33.4%	83%

DISCUSSION

The major goal of breast screening is to detect a high percentage of cancers which is reflected as the cancer detection rate. Besides, it is intended to find the cancers within an acceptable rate of recommendation for additional imaging and recommendation for tissue diagnosis. These rates are calculated as recall rate and positive predictive values. Another goal of screening is to detect a high percentage of small (minimal cancer), node-negative and early-stage cancers. Auditing gives a quantifiable result to monitor these goals. This study showed that detection of a desirable rate of minimal cancers and early-stage cancers with an appropriate recall rate and PPV is possible in a clinical setting where screening is held.

Although there is an official national breast cancer screening programme, there is still not a national mammographic registry system that will allow us to reach a reliable source to track the cancer diagnosis of all our cases in the first year after screening. This data is needed to identify the real negative and false-positive screening mammograms which will lead to the calculation of reliable sensitivity, specificity and negative predictive values. However, few audits in the literature on breast cancer screening audits in Turkey exist (10,11,15). These studies have encountered similar problems in retrieving and mining their data. On the other hand, the follow-up of all women with negative or benign findings was not possible to reach real true negative values. We have calculated the negative predictive value (99.88%) and sensitivity (90%) depending on the negative or benign findings in US examinations which covers the 85.9% of all our screening mammograms. Although this calculation is not ideal, we believe that it still can reflect the quality of the mammography service.

Our cancer detection rate with mammography (0.74%) was slightly higher than the previous mammography screening audits (22% than Tunçbilek's, 9% than Turk's, 29% than Kayhan's studies) previously published in Turkey but was in the line of the recommendation of ACR (≥0.25%). Cancer detection rate reflects the quality of mammographic practice. Our higher rate may reflect the impact of tomosynthesis. Recent tomosynthesis studies showed an improvement in cancer detection by at least 30% (12,13).

In the mammographic screening of women with no personal history of breast cancer our PPV1 (positive cases) which reflects perceptual skills and PPV2 (biopsies recommended) are within the normal range compared to ACR recommendations. On the other hand, we reached

a higher PPV3 (cancer yield of biopsies) value of 52.9%. However, 34.6% of our patients to whom biopsy was recommended were lost during follow-up. We believe that this is the main reason for our high PPV3 rate.

Recall rate is the percentage of patients recommended for further imaging. One of the aims of mammography is to detect cancers with a low recall rate. Our recall rate (11%) is within the recommended limits by ACR (ACR recommendation: 5-12%).

Major prognostic factors in breast cancer are size, stage and lymph node status. The ability to detect cancers at a curable and an early stage is reflected by the percentage of minimal cancers and lymph node negativity. Our percentage of minimal cancers (%50) is within the limits and shows the impact of effective screening. The percentage of stage 0-1 cancers (%77) are above the desired goals. However, lymph node negativity for invasive cancers (%64) was slightly lower than the normal range while it was 77% for both invasive and insitu cancers. Although our early-stage cancer rates are above the benchmarks, low lymph node negativity is a contradiction. This can be explained by our low number of screening patients and detected invasive cancers.

The women with dense breast ratios were high in our study group with 58.94% of women having BI-RADS type C and D breast density. This rate was higher than the Turkish data from the national screening program which showed 38.5% (C type 34%, D type 4.5%) dense breast (14). However, it was closer to the findings of an organized screening program in a county of Istanbul which was 52.7% (C type 36.8%, D type 15.9%) (15). This similarity shows that the density might be higher in Istanbul compared to the women in the country in general.

Limitations of this audit are: First, the screening performance and cancer detection outcomes were calculated based on the raw data that was documented manually. However, difficulties in the follow up of women that were lost after a positive or suspicious finding were one of the main obstacles in reaching perfect PPVs. Patients were lost after a positive screening mammogram finding. They change their doctor for a second opinion or prefer to go to another clinic according to a recommendation. Due to the lack of a national reliable cancer registry program, it is not possible to reach a new cancer diagnosis of the screened women without positive findings. Second, we manually collected the raw data from the questionnaires

of the patients and we lost the data of 9.76% of the patients. Difficulties related to manual data collection methods and lack of computer-based data registry and mining are important and a common problem for countries with limited resources like Turkey (16). Dedicated data collection and analysis methods for auditing will enable fast and easy reporting. Also, standardized reporting and regular documentation of these reports are the essentials for a healthy auditing. "Third, although we have recorded every abnormal mammogram which needed an additional diagnostic study as BI-RADS 0, in most cases we have completed the examination with the additional imaging on the same day of the screening. Furthermore, in order to simulate a mammogram only screening, BI-RADS 0 was used independently and separately from the simultaneous screening US findings, even in cases where the screening US enlightened the mammography finding. The readers might be biased as they were not blinded to the US findings and this creates another limitation in the study."

We believe that every breast imaging center must audit its practice to evaluate its effectiveness mainly in screening performance and it can be extended to diagnostic and cancer follow-up examinations. Auditing data is necessary for the radiologists to understand the quality of their daily work and robust feedback in presenting the needs for further improvement.

Acknowledgement

We thank Ayşe Yazıcı, Hülya Güç, Tuğçe Hirik, Ceylan Akdur, Kevser Başak, Gözde Alpay and Sennur Karagöz for their contributions.

REFERENCES

- 1. Acheson HWK. Medical Audit and General Practice. Vol. 1, Lancet. 1975. 511–513 p. Crossref
- 2.CJ. D. 2013 ACR BI-RADS Atlas: Breast Imaging Reporting and Data System, 5th edition. Am Coll Radiol. 2014;537–53.
- 3.A Sickles E. Quality Assurance: How to audit your own mammography practice. Vol. 30, Radiologic clinics of North America. 1992. 265–275 p.
- 4. Gürdemir B, Aribal E. Assessment of mammography quality in İstanbul. Diagnostic Interv Radiol. 2012; 18: 468-472. Crossref
- 5.Sickles EA, Miglioretti DL, Ballard-Barbash R, Geller BM, Leung JWT, Rosenberg RD, et al. Performance benchmarks for diagnostic mammography. Radiology. 2005 Jun;235(3):775–90.
- 6.Feig SA. Auditing and benchmarks in screening and diagnostic mammography. Radiol Clin North Am. 2007 Sep;45(5):791–800, vi. Crossref
- 7.Carney PA, Sickles EA, Monsees BS, Bassett LW, Brenner RJ, Feig SA, et al. Identifying minimally acceptable interpretive performance criteria for screening mammography. Radiology 2010 May; 255(2):354–61. Crossref

- 8.Giuliano AE, Connolly JL, Edge SB, Mittendorf EA, Rugo HS, Solin LJ, et al. Breast CancerMajor changes in the American Joint Committee on Cancer eighth edition cancer staging manual. CA Cancer J Clin. 2017 Jul;67(4):290–303. Crossref
- 9.Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Vol. 17, Annals of surgical oncology. United States; 2010. p. 1471–4. Crossref
- Türk A, Aribal E. Medical audit of mammography screening examinations. Marmara Med J. 2002; 15: 27-31. Crossref
- Tuncbilek I, Ozdemir A, Gultekin S, Ogur T, Erman R, Yuce C. Clinical outcome assessment in mammography: an audit of 7,506 screening and diagnostic mammography examinations. Diagn Interv Radiol. 2007 Dec;13(4):183–7.
- Zackrisson S, Lang K, Rosso A, Johnson K, Dustler M, Fornvik D, et al. One-view breast tomosynthesis versus two-view mammography in the Malmo Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. Lancet Oncol. 2018 Nov;19(11):1493–503. Crossref
- Skaane P, Sebuodegard S, Bandos AI, Gur D, Osteras BH, Gullien R, et al. Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective populationbased Oslo Tomosynthesis Screening Trial. Breast Cancer Res Treat. 2018 Jun;169(3):489–96. Crossref
- Gultekin M, Ozturk C, Ozturker Z, Boztaş G, Hatice Turan S, Dundar S, et al. Centralization of mammography reporting with mobile trucks: Turkish experience. Vol. 10, Preventive Medicine Reports. 2018.
- Kayhan A, Gurdal SO, Ozaydin N, Cabioglu N, Ozturk E, Ozcinar B, et al. Successful first round results of a Turkish breast cancer screening program with mammography in Bahcesehir, Istanbul. Asian Pac J Cancer Prev. 2014;15(4):1693–7. Crossref
- Aribal E, Mora P, Chaturvedi AK, Hertl K, Davidovic J, Salama DH, et al. Improvement of early detection of breast cancer through collaborative multi-country efforts: Observational clinical study. Eur J Radiol. 2019 Jun;115: 31–8. Crossref