Analysis of diagnostic methods of a population-based breast cancer early detection screening programme

Katarzyna Szymoniak1,CD, Maria Pytel1,AB, Witold Malinowski1A, Wojciech Pytel1AB, Dorota Fryc1AB, Małgorzata SzkupSEF

1 Department of Obstetrics and Pathology of Pregnancy, Faculty of Health Sciences, Pomeranian Medical University, Szczecin, Poland
2 Salve Medica, Health Care Unit Ltd., Poland
3 Faculty of Health Sciences, Masovian Public University, Płock, Poland
4 Barska Diagnostics and Treatment Centre Ltd., Poland
5 Department of Nursing, Faculty of Health Sciences, Pomeranian Medical University, Szczecin, Poland

A – Research concept and design, B – Collection and/or assembly of data, C – Data analysis and interpretation, D – Writing the article, E – Critical revision of the article, F – Final approval of article

Key words
breast cancer, mammography, sonomammography, screening programme

Abstract
Introduction and objective. Breast cancer is the most common malignant tumour in women in Poland. In 2006, the Population-based Breast Cancer Early Detection Screening Programme was introduced in Poland to decrease the mortality of women due to this cancer. This study aimed to analyse the diagnostic methods used in the Population-based Breast Cancer Early Detection Screening Programme in the experience of one facility in Poland.

Materials and method. The material for the study consisted of 1,411 questionnaires from women aged 50–69 qualified for enhanced diagnostics by a mammography performed in the first stage of the programme. During the study, a retrospective analysis of the documentation was performed.

Results. Analysis of the convergence of BI-RADS scores between MMG and USG showed that the highest convergence was confirmed for a BI-RADS score of 4 and the lowest for a BI-RADS score of 5. Comparing the sensitivity and specificity of MMG and USG after the acquisition of BI-RADS, BAC were more frequently chosen for enhanced diagnostics of breast cancer. However, cancer was significantly more often confirmed by BAG. Tumours were biopsied more frequently for BI-RADS scores of 4 and 5 in both MMG and USG. Core-needle biopsy was more frequently used to diagnose BI-RADS scores of 3, 0 and 4 in mammography.

Conclusions. The obtained results showed the highest sensitivity of breast cancer screening methods for BI-RADS MMG (4+5) and ultrasound (3+4+5). The highest convergence of results between MMG and USG is shown for a BI-RADS score of 4. Breast tumours are more often diagnosed by BAC, but cancer is more often confirmed by BAG.
INTRODUCTION

Breast cancer is the most common malignant tumour in women in Poland. The basic examination in the diagnosis of breast cancer is mammography. In 2006, the Population-based Breast Cancer Early Detection Screening Programme was introduced in Poland to decrease up to 40% the mortality of women due to this cancer [1]. The screening programme comprises two stages: first stage – mammography, second stage – ultrasonography and additional fine-needle or core-needle biopsy.

Screening mammography aims to detect focal breast lesions that are impalpable during a normal examination. If the lesion is suspicious, the woman should be referred for the second stage of the programme which is an ultrasound examination that helps to estimate the malignancy of the lesion, and to qualify the woman for enhanced diagnostics (needle biopsy). It is recommended that lesions suspected of being malignant are biopsied using fine-needle or core-needle biopsy, which allows cytological and histopathological verification of the lesions, which is crucial in planning the patient’s treatment.

In Poland, breast screening is performed every 2 years in women aged 50–69 [2]. The European Society of Breast Imaging (EUSOBI) and national radiological societies recommend mammography as part of a population-based screening. Despite the inherent limitations associated with its sensitivity and specificity, mammography remains the primary tool for population-wide screenings. However, the practical implementation of the Population-based Breast Cancer Early Detection Screening Programme and the improvement of screening procedures is the responsibility of screening centres and requires a careful analysis of the measures that qualify women for further enhanced diagnostics with the lowest possible diagnostic error.

This study is aimed at analysing the diagnostic methods used in the Population-based Breast Cancer Early Detection Screening Programme in the experience of one facility in Poland.

MATERIALS AND METHOD

The material for the study consisted of 1,411 questionnaires from women aged 50–69 and participating in the second stage of the Population-based Breast Cancer Early Detection Screening Programme. Mammography, performed on the participants in the first stage of the programme, was described according to the Breast Imaging-Reporting and Data System as a lesion requiring enhanced diagnostics. This is a system designed to standardize the descriptions of mammography, ultrasound, and MRI, and includes the following categories of lesion description: 0 – final assessment incomplete, 1 – normal, 2 – benign lesion, 3 – probably benign lesion, 4 – suspicious lesion, 5 – lesion with high probability of malignancy.

Medical records of 749 women from 2010 and 662 women from 2013 were used in the study because these years saw the highest number of patients with inconclusive BI-RADS MMG results, i.e. categories 0, 3, 4, 5. Women with BI-RADS MMG 1 and 2, who did not require a further diagnosis for breast cancer, were excluded from the study. The studied women were divided into 3 age groups: group I – women aged 50–55, group II – women aged 56–62, group III – women aged 63–69.

The material was obtained from Salve Medica Health Centre in Łódź, which participated in the first and second stage of the Population-based Breast Cancer Early Detection Screening Programme. During the study, a retrospective analysis of the documentation was performed. The Salve Medica Health Centre was selected for the study due to the high enrolment of patients in the first and second stage of the programme. In addition, this unit offers complete diagnostics, treatment, and rehabilitation in breast cancer. During the study, a retrospective analysis of the following documentation used in the programme was conducted: ‘Referral to the Population-based Breast Cancer Early Detection Screening Programme’, ‘Supplementary Examination Cards of the Population-based Breast Cancer Early Detection Screening Programme’, and the results of microscopic examinations of the material collected from the breast gland by means of fine- and core-needle biopsies. The participants’ sensitive data included in the analysed documentation were prepared for further analysis by the implementers of the programme in such a way that in all documents the name and surname, date of birth, Personal Identification Number and contact information were blanked. Consecutive code numbers were assigned to the patients, taking into account the year in which the examination was performed (2010 or 2013) and the woman’s age at the time of her examination. Written consent was obtained from the CEO of the Salve Medica Health Centre to release the records for this analysis. The study was approved by the Bioethics Committee of the Pomeranian Medical University in Szczecin (KB-0012/253/06/18).

Yates’ chi-square test and Pearson’s chi-square test were used to compare qualitative variables. Sensitivity, specificity, and the AUC (Area Under Curve) index were estimated to assess the diagnostic value of the research methods. A two-sided significance level of p<0.05 was assumed for all the tests used. Analysis was performed using the Statistica 12.5 software.

RESULTS

In the Salve Medica Health Centre, 1,411 women participated in the second stage of the Population-based Breast Cancer Early Detection Screening Programme. Among them, there were 296 fine-needle biopsies and 120 core-needle biopsies, identifying 149 breast tumours. The patients most frequently presented with mammography BI-RADS scores of 0 (678 women) and 3 (512 women), followed by BI-RADS scores of 4 (171 women) and 5 (47 women). Enhanced diagnostics most frequently showed ultrasound BI-RADS scores of 1 and 2 (972 women), and 4 (243 women), followed by BI-RADS scores of 3 (180 women) and 5 (16 women).

Of all the eligible women, 53% were enrolled in 2010 and 47% in 2013. Their median age was 58±5.4 years. Women from age group II, i.e. aged 56–62 (38.91%), enrolled most frequently for the second stage of the programme, while the oldest women from age group III (24.24%) enrolled least frequently. Group I (aged 50–55) consisted of 352 women (36.85%).

Analysis of mammography results qualifying for enhanced diagnostics showed that BI-RADS score of 0, i.e. an incomplete test, was predominant in the youngest group, i.e. women at perimenopausal age, whereas lesions described in BI-RADS scores of 3, 4, 5 were more frequently diagnosed in women over 56 years of age.
In the group of patients with mammography BI-RADS score of 0, a total of 101 fine-needle biopsies and 16 core-needle biopsies were performed, diagnosing breast cancer in 17 cases. For ultrasound BI-RADS score of 2, there was 1 case, for ultrasound BI-RADS score of 3, there were 2 cases, and for ultrasound BI-RADS score of 4, there were 14 cases of cancer. A BI-RADS score of 5 did not occur in ultrasound scans.

Convergence results between mammography and ultrasound scan in the diagnosis of breast cancer. An important aspect of the study was analysis of the convergence of BI-RADS scores between mammography and ultrasound. Analysis of the results of the women examined showed that the greatest convergence of the results was confirmed for ‘suspicious lesions’, described in both studies as a BI-RADS score of 4 (49.71%). Cohen’s kappa coefficient was 0.198 (Tab. 1). Similar results were obtained by dividing the women into age groups. In the first group, 45.61% of breast lesions were described by this score, whereas in the oldest group, 40.91% of breast lesions were described by a BI-RADS score of 4 (Tab. 2).

Interestingly, the lowest convergence of mammography and ultrasound findings was confirmed for lesions with a high probability of malignancy (BI-RADS score of 5) both for individual age groups and in total, without separating age ranges (14.00%) (Tab. 1, Tab. 2). Analysis of Table 2 regarding the prevalence of breast lesions requiring additional diagnostic procedures revealed that abnormal breast lesions were more frequently detected during mammography in women aged between 56–62 than in the other groups of women. The second group in terms of the prevalence of suspicious lesions were younger patients, aged less than 55, and the least abnormal results were obtained from patients aged over 62 years.

To compare the sensitivity and specificity of mammography and ultrasound, a model was created in which BI-RADS scores of 4 and 5 in mammography, i.e. suspicious lesions and high probability of malignancy, and BI-RADS scores of 3, 4 and 5 in ultrasonography, i.e. uncertain, suspicious lesions, high probability of malignancy, were combined. This showed that the BI-RADS MMG (4+5) and the BI-RADS MMG (4+5), and ultrasound (3+4+5) model have the highest sensitivity in detecting pathological changes in the breast, while the specificity of these models is lower than BI-RADS MMG (4+5) and ultrasound (4+5) model. The difference was not statistically significant (p>0.05) (Tab. 3).

Table 1. Convergence of mammography and ultrasound results for different BI-RADS scores, without age grouping

| Type of test | USG BI-RADS 1 | USG BI-RADS 2 | USG BI-RADS 3 | USG BI-RADS 4 | USG BI-RADS 5 | Total | Cohen’s kappa |
|--------------|---------------|---------------|---------------|---------------|---------------|-------|--------------|
| MMG BI-RADS 3 | 187           | 173           | 89            | 62            | 1             | 512   | 0.198        |
| MMG BI-RADS 4 | 42            | 19            | 17            | 85            | 8             | 171   |              |
| MMG BI-RADS 4 | 24.56%        | 11.11%        | 9.94%         | 49.71%        | 4.68%         | 50    |              |

Table 2. Convergence of mammography and ultrasound results BI-RADS scores for different age groups

| Age       | Type of test | USG BI-RADS 1 | USG BI-RADS 2 | USG BI-RADS 3 | USG BI-RADS 4 | USG BI-RADS 5 | Total | Cohen’s kappa |
|-----------|--------------|---------------|---------------|---------------|---------------|---------------|-------|--------------|
| 50-55     | MMG BI-RADS 3 | 57            | 82            | 24            | 21            | 0             | 184   | 0.00%        |
| 50-55     | MMG BI-RADS 4 | 30.98%        | 44.57%        | 13.04%        | 11.41%        | 0.00%         | 100%  |              |
| 50-55     | MMG BI-RADS 5 | 22.81%        | 14.04%        | 8.77%         | 45.61%        | 8.77%         | 100%  |              |
| 56-62     | MMG BI-RADS 3 | 77            | 56            | 35            | 28            | 1             | 197   |              |
| 56-62     | MMG BI-RADS 4 | 39.09%        | 28.43%        | 17.77%        | 14.21%        | 0.51%         | 100%  |              |
| 56-62     | MMG BI-RADS 5 | 22.86%        | 10.00%        | 5.71%         | 58.57%        | 2.86%         | 100%  |              |
| 63-69     | MMG BI-RADS 3 | 0             | 0             | 0             | 0             | 0             | 0     |              |
| 63-69     | MMG BI-RADS 4 | 40.46%        | 26.72%        | 22.90%        | 9.92%         | 0.00%         | 100%  |              |
| 63-69     | MMG BI-RADS 5 | 29.55%        | 9.09%         | 18.18%        | 40.91%        | 2.27%         | 100%  |              |
BI-RADS scores of 3, 4 and 5 were an indication for biopsy of a breast lesion. In the group of 1,411 patients, 416 women were referred for biopsy (71.15% – BAC; 28.85% – BAG) and 149 (35.81%) and breast tumours were detected. Analysis of the course of enhanced diagnostics in age groups showed that in the youngest group, aged 50–55, biopsies were performed in 29.81% of the women in whom cancer was diagnosed in 6.73%; in women aged 56–62, biopsies were performed in 30.24% of the examined and cancer confirmed in 13.30%; in the oldest group, aged 63–69, biopsies were performed in 27.78% of the cases, and cancer confirmed in 11.99%.

Statistical analysis showed that breast cancer was significantly more frequently diagnosed in the biopsy material in women aged 56–69. Statistical significance was taken at p<0.001 (Tab. 4).

Table 4. Confirmation of breast cancer in biopsies by age group

| Age        | No. of patients | No. of biopsies | % of patients with biopsy | No. of cancer confirmations | % of cancer confirmations | p   |
|------------|----------------|----------------|--------------------------|----------------------------|--------------------------|-----|
| 50-55      | 520            | 155            | 29.81%                   | 35                         | 6.73%                    |     |
| 56-62      | 549            | 166            | 30.24%                   | 73                         | 13.30%                   | 0.001|
| 63-69      | 342            | 95             | 27.78%                   | 41                         | 11.99%                   |     |

Fine-needle biopsy (71.15%) was more frequently chosen for enhanced diagnostics of breast cancer than core-needle biopsy (28.85%). In contrast, from all detected cancers, 92.50% cancer was confirmed by core-needle biopsies (p<0.001). At the same time, fine-needle biopsies were significantly less frequently used in the diagnosis of women aged 63–69, and significantly more frequently used in the diagnosis of women aged 50–62 (p=0.04).

Breast tumours were biopsied (fine-needle biopsy or core-needle biopsy) significantly more often in case of both mammography and ultrasonography BI-RADS scores of 4 and 5. Furthermore, core-needle biopsies (83.0%) were significantly more frequently used than fine-needle biopsies (17.0%) to diagnose BI-RADS 5 tumours in both mammography and ultrasound. In contrast, fine-needle biopsies were significantly more frequently used to diagnose lesions with mammography BI-RADS scores of 3, 0 and 4 (92.0%; 84.3%; 53.6%) (Tab. 5).

**DISCUSSION**

Screening mammography has been recognized worldwide as an effective tool in reducing mortality from breast cancer among women [1, 2]. In Poland, there has been a gradual decrease in the standardized death rate from this disease since 2007. Unfortunately, the decrease is slower than in most European Union countries [3]. These facts encourage the search for procedures that increase the effectiveness of diagnosing breast cancer early within the framework of the Population-based Breast Cancer Early Detection Screening Programme. One way to do this is to analyze the elements of the programme that affect its quality and effectiveness.

The presented study showed that through screening mammography, 1,411 women were qualified for enhanced diagnostics in 2010 and 2013. As a result of the examinations, 972 benign lesions were diagnosed and 439 lesions were identified on the spectrum of ‘probably benign’ to ‘probably malignant’.

Breast ultrasonography is an obligatory examination performed in the second stage of the programme and its primary aim is to clarify the nature of breast lesions detected by screening mammography. In both of these imaging examinations, the detected lesion is classified by a BI-RADS score [4]. The use of the BI-RADS classification allows for consistent interpretation of the lesion, comparison and recommendations, and improvement of breast imaging diagnostic methods.

An important aspect of the current study was to assess the convergence of mammography and ultrasonography results based on the BI-RADS classification. The study also showed discordance in the BI-RADS scores between mammography and ultrasonography for all the women, without division into years. For a BI-RADS score of 3 it was 82.60% and for a BI-RADS score of 5 it was 86.00%, whereas the highest concordance was confirmed for a BI-RADS score of 4–50.20%. Similar results were shown by Stavros et al. [5] in their study of breast lesions classified as BI-RADS 3 (79%), while different results were obtained for a BI-RADS score of 4, where the discordance between mammography and ultrasound was as high as 81%, and for a BI-RADS score of 5 – only 13%. Also, Kuhl et al. [6] showed discordance of results between mammography (9.5%) and ultrasound (16.7%) in BI-RADS 3, which was statistically significant. The demonstrated lack of convergence between the results in the BI-RADS 3 category in a similar value in our study and the literature is a desirable phenomenon for the assumed diagnostic goal, which is a definite change in the classification of benign or malignant lesions by ultrasound. The significant difference in the convergence of mammography and ultrasound results for BI-RADS scores of 4 and 5 in the literature and this study are probably due to the use of non-uniform, conservative and subjective criteria for qualifying breast tumours to particular BI-RADS categories, both in mammography and in ultrasound performed in the Population-based Breast Cancer Early Detection Screening Programme. A study by Hodorowicz-Zaniewska et al. [7] showed no difference
in the BI-RADS scores obtained from mammography and ultrasound because some lesions were visible only in one of the imaging methods. Therefore, the highest mammography or ultrasound BI-RADS score was assumed as a combined BI-RADS score. In contrast, Kuhl et al. [6] showed that out of 40 breast cancers diagnosed by imaging, only 21 were diagnosed after combined mammography and ultrasonography. It should be noted that studies on the convergence results between mammography and ultrasound are considered in few literature items.

The sensitivity and specificity of the diagnostic methods used in detecting breast cancer is extremely important. In the presented study, mammography BI-RADS scores of 4 and 5 and ultrasound BI-RADS scores of 3, 4 and 5 were combined to determine the highest sensitivity and specificity of the breast imaging methods. It was shown that the sensitivity of the examination was at a similar level in the BI-RADS MMG (4+5) and the BI-RADS MMG (4+5), and ultrasound (3+4+5) model. Whereas the highest specificity of the diagnostic methods was demonstrated for the BI-RADS MMG (4+5) and ultrasound (4+5) model. A similar study was conducted by Dobruch-Sobczak [8] which combined the ultrasound BI-RADS scores. The best model was a combination of ultrasound BIRADS score of 4 and 5, in which the sensitivity was 76.9% and the specificity 96.67%. Different results were presented by Jassem et al. [9], which showed that the sensitivity of mammography was 80%; however, in the study by Kuhl et al. [6], mammography had the lowest sensitivity (25%) and specificity at 96.8%.

An absolute confirmation of malignant breast cancer that entitles the patient to oncological treatment is the detection of cancer cells in the sampled tumour tissue. The current study shows that out of all the examined patients, 29.48% were referred for biopsy and 35.81% were diagnosed with breast cancers. A similar percentage of patients referred for biopsies after breast ultrasonography (21%) was demonstrated in the study by Dobruch-Sobczak [8] in which 39% of patients were diagnosed with cancer. In the current study, breast cancer was significantly more frequently diagnosed in the biopsy material in women aged 56-69 (p<0.001). Hodorowicz-Zaniewska et al. [7] also showed that the mean age of women with malignant breast lesions was 56.8±13.02. On the other hand, the results of the study by Dobruch-Sobczak [8] differed from the above – the mean age of patients with diagnosed breast cancer was lower and amounted to 55.07.

In the current study, fine-needle biopsy was chosen more often (71.15%) than core-needle biopsy (28.85%) for enhanced diagnostics of breast cancer; however, cancer was significantly more frequently confirmed by core-needle biopsies (92.50%). Similar results were presented by Łukasiewicz et al [10] which showed that 82% of breast cancers were diagnosed using core-needle biopsies, and that the sensitivity and specificity of fine-needle biopsies varied and was lower than in core-needle biopsies in breast cancer diagnostics. However, Aker et al. [11] proved that fine-needle biopsies and core-needle biopsies are equivalent diagnostic techniques, especially for suspicious breast lesions.

In this study, breast tumours were biopsied (fine-needle biopsy or core-needle biopsy) significantly more often as a result of their classification in mammography and ultrasonography. BI-RADS scores 4 and 5. Importantly, core-needle biopsies were used significantly more often than fine-needle biopsies to diagnose tumours with a BI-RADS score of 5 in both mammography and ultrasonography. Fine-needle biopsies were significantly more frequently used to diagnose mammography BI-RADS 0, 3 and 4 lesions. Similar results were presented by Dobruch-Sobczak [8] in which breast lesions with a BI-RADS score of 3 were diagnosed by fine-needle biopsies, while lesions with a BI-RADS score of 4 were diagnosed based on core-needle biopsies. Bednarski et al. [4] also showed that lesions with a BI-RADS score of 3 should be diagnosed by fine-needle biopsies, or checked in 6 months by ultrasound.

**Limitations of the study.** First of all, cross-sectional data from 2010 and 2013 was used in the study which may suggest that the findings are outdated. Secondly, the study group consisted of women from a single facility in Poland; therefore, caution should be exercised when interpreting the entire population of women. Thirdly, it was not possible to present the most up-to-date results of other authors’ studies because there are no new reports on the analysis of diagnostic methods in breast cancer screening. Further research should address a similar analysis including magnetic resonance imaging among diagnostic methods, and an analysis of several facilities in Poland in different years. Despite these limitations, the study has several advantages. Firstly, it analyses the convergence of BI-RADS mammography and ultrasonography results, which allows assessment of the effectiveness of a diagnostic method in breast cancer screening. Secondly, it evaluates the sensitivity and specificity of mammography and ultrasonography in a combination of BI-RADS MMG and ultrasound, which suggests the best model for breast lesion imaging. Thirdly, it analyses the use of fine needle and core needle biopsy in relation to BI-RADS MMG and ultrasound category tumours.

**CONCLUSIONS**

1. The obtained results showed the highest sensitivity of breast cancer screening methods for BI-RADS MMG (4+5) and ultrasound (3+4+5).
2. The highest convergence of results between mammography and ultrasound is shown for BI-RADS score of 4.
3. Breast tumours are more often diagnosed by fine-needle biopsy, but cancer is more often confirmed by core-needle biopsy.

**REFERENCES**

1. Sardanelli F, Aese HS, Álvarez M, et al. Position paper on screening for breast cancer by the European Society of Breast Imaging (EUSOBI) and 30 national breast radiology bodies from Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Israel, Lithuania, Moldova, the Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Spain, Sweden, Switzerland and Turkey. Eur Radiol. 2017; 27(7):2737–2743.
2. Wypij N. Magnetic resonance imaging in breast cancer – indications. Lett Oncol Sci. 2017; 14(4):110–115.
3. Krajowy rejestr nowotworów. MJP Online Early: http://onkologia.org.pl/ (access: 2020.09.24).
4. Bednarski P, Dobruch-Sobczak K, Chrupowicki E, et al. Breast ultrasound scans – surgeons’ expectations. J Ultrasound. 2015; 15(61):164–171.
5. Stavros AT. Ultrasonografia piersi. Warsaw: Medipage; 2007.
6. Kuhl CK, Schrading S, Leutner CC, et al. Mamnography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer. J Clin Oncol. 2005; 23(33): 8469–8476.
7. Hodorowicz-Zaniewska D, Brzuszkiewicz K, Szpor J, et al. Clinical predictors of malignancy in patients diagnosed with atypical ductal hyperplasia on vacuum-assisted core needle biopsy. Videosurgery Miniinv. 2018; 13(2): 184–191.
8. Dobruch-Sobczak K. The differentiation of the character of solid lesions in the breast in the compression sonoelastography. Part I: The diagnostic value of the ultrasound B-mode imaging in the differentiation diagnostics of solid, focal lesions in the breast in relation to the pathomorphological verification. J Ultrason. 2012; 12(51): 402–419.
9. Jassem J, Krzakowski M. Breast cancer. Oncol Clin Pract. 2018; 14(4): 171–215.
10. Łukasiewicz E, Zierniecka A, Jakubowski W, et al. Fine-needle versus core-needle biopsy – which one to choose in preoperative assessment of focal lesions in the breasts? Literature review. J Ultrason. 2017; 17(71): 267–274.
11. Aker F, Gümrükçü G, Onomay BC, et al. Accuracy of fine-needle aspiration cytology in the diagnosis of breast cancer a single-center retrospective study from Turkey with cytohistological correlation in 733 cases. Diagn Cytopathol. 2015; 43(12): 978–986.