

Original paper

## Evaluation of the positive predictive value (PPV3) of ACR BI-RADS category 4 and 5 based on the outcomes of Invasive Diagnostic Office in an outpatient clinic

Krzysztof Koziłek<sup>1A,B,C,D,E,F</sup>, Natalia Stranz-Walczak<sup>2B,C,E,F</sup>, Paweł Gajdzis<sup>1B</sup>, Katarzyna Karmelita-Katulska<sup>2E</sup>

<sup>1</sup>Poznan Center of Cancer Prevention and Epidemiology in memory of A. Pienkowska, Poznan, Poland

<sup>2</sup>Department of General Radiology and Neuroradiology, Poznan University of Medical Sciences, Poland

### Abstract

**Purpose:** The American College of Radiology (ACR) BI-RADS classification is the applicable for breast lesion assessment. BI-RADS categories 4 and 5 need to be followed by biopsy. The aim of our study was to evaluate the tissue biopsy-proven positive predictive value (PPV3) for BI-RADS 4 (and its subcategories) and for BI-RADS 5, and BI-RADS distribution, in comparison to ACR assumptions and literature.

**Material and methods:** We retrospectively analysed biopsies performed in our outpatient clinic in 2017. Our target group of patients consisted of 797 patients at the average age of 52 years.

**Results:** BI-RADS 5 constituted 12% of cases (95 cases), and BI-RADS 4 88% (698 cases). Within BI-RADS 4 subdivisions there were 359 cases in 4a (45.3%), 215 in 4b (27%), and 124 in 4c (15.6%). Overall PPV3 was 34.8%. BI-RADS 5 category PPV3 number was 97.89%. In category BI-RADS 4 the values of PPV3 equalled 26.22% without subdivision and 3.6%, 27.9%, and 88.7% for subcategories BI-RADS 4a, 4b, and 4c, respectively.

**Conclusions:** BI-RADS categorisation by radiologists in the studied group matches the literature data according to achieved PPV and BI-RADS percentage distribution. The stratification of cancer risk among categories was proven with Mann-Whitney *U* test *p* value < 0.005. There was a statistically important unaccepted difference of PPV3 between core biopsy and vacuum-assisted biopsy, which needs further investigation.

**Key words:** breast cancer, breast biopsy, BI-RADS classification.

### Introduction

The American College of Radiology (ACR) BI-RADS classification comprises combined lexicon, result, and conclusion of breast examination. The classification was first introduced in 1993, seven years after its proposal, as a unified reporting for mammography examination. In 2003 it was adapted to breast ultrasonography and magnetic resonance, and additionally, due to cancer risk of biopsy outcome, BI-RADS 4 was divided into three subcategories:

4a, 4b, and 4c [1,2]. The recent fifth edition of the lexicon was published in 2013 [3]. BI-RADS category 5 is reserved for cancer-like lesions on diagnostic imaging. BI-RADS 4 includes suspicious lesions with some likelihood of malignancy [3,4]. For both BI-RADS categories 4 and 5 lesions biopsy is mandatory, which makes them suitable for the calculation of biopsy-proven positive predictive value (PPV) [5,6].

In the United States of America, mammography BI-RADS is under federal law and control (FDA's MMG Quality

### Correspondence address:

Krzysztof Koziłek, Poznan Center of Cancer Prevention and Epidemiology in memory of A. Pienkowska, Poznan, 24/26 Kazimierza Wielkiego, 61-863 Poznań, Poland, e-mail: [k.kozielek@gmail.com](mailto:k.kozielek@gmail.com)

### Authors' contribution:

A Study design · B Data collection · C Statistical analysis · D Data interpretation · E Manuscript preparation · F Literature search · G Funds collection

Standards Final Rule), but BI-RADS for ultrasonography is not subject to this regulation. Furthermore, the Food and Drug Administration (FDA) rule makes the BI-RADS 4 subcategories non-obligatory.

In Europe, besides BI-RADS, there is the coexisting R1-5 score system proposed by the European Society of Breast Cancer Specialists (EUSOMA), but this has never gained popularity in Poland [7]. The Polish national breast cancer screening program is based on the BI-RADS lexicon but without the BI-RADS 4 subdivision. There is no legal implementation of BI-RADS classification for diagnostic purposes, but the scale is well known and commonly used among performers of breast diagnostics, and it is recommended [8]. According to the data from the National Mammography Database in the United States of America, about one third of BI-RADS 4 cases were reported in subcategories [9]. From our experience this does not apply to Poland, because even if there is only BI-RADS 4 in screening outcome, sustained on recall visit, it is to be subcategorised.

The classification seems to be free from the influence of patient factors like age or family history of breast cancer [10]. Therefore, the number of features linked to a performing physician, influence BI-RADS categorisation performance, among them the experience factor is the most important, and highlights the differences among groups of years of experience and the number of mammograms assessed yearly [11,12].

Histopathological assessment is the most reliable verification of the BI-RADS categorisation. Besides confirmation or refutation of cancer, it provides data such as histopathological type, cancer cell differentiation, and heterogeneity of tumour [13].

There is international agreement that surgical biopsy is a last-hope procedure and should not be commonly performed [7,14]. Breast tissue biopsy procedures have proven their accuracy and have effectively replaced the surgical (open) biopsy for histopathology.

There are two types of tissue breast biopsies: core biopsy (CB) and vacuum-assisted biopsy (VAB). Both CB and VAB are characterised by a very high negative predictive value, reaching as much as 99.1% [15].

In Poland, applicable guidelines for CB have been approved by the cooperating societies of pathologists, oncologists, surgeons, and radiologists. VAB indications have been precisely defined as part of the Polish Ministry of Health Guidelines (third attachment of the list of guaranteed services of outpatient treatment procedures and their implementation conditions) [13,16]. Because Polish biopsy indication correlates with ACR diagnostic recommendations and BI-RADS classification is commonly used, we decided to evaluate BI-RADS categorisation from the perspective of our biopsy office, regardless of the basis of the diagnosis – whether it was based on mammography or ultrasonography performed together or separately.

## Material and methods

The outcomes of all breast biopsies performed in 2017 in the Invasive Diagnostic Office of the diagnostic outpatient clinic were the subject of our retrospective analysis. The office offers its services to the Centre's needs, i.e. screening, oncology, and oncological surgery outpatient clinics and within agreements with other health centres, including commercial biopsies. CB and VAB biopsies in the examined cases were performed under the control of an ultrasound examination.

In 2017 there were 895 biopsies performed. Exclusion criteria were the lack of pre-biopsy examination outcome, result without BI-RADS category, multiple biopsies of the same breast at the same time in multiple foci cancer occurrence, and axillary lymph nodes biopsies. The biopsies of breast lesions that proved to be lymph nodes on histopathological examination were included. Finally, there were 797 cases in the studied group.

The group consisted of 795 women and two men with the average age of 52 years, ranging from 18 to 92 years.

There were 236 VAB and 561 CB performed. For the CB, 14G 2 cm wide window needle with an automated biopsy gun was used, whereas VAB was performed with an Encore system VS3000 with 7G, 10G, and 12G needles. Both CB and VAB biopsies were performed under the control of a Hitachi HI VISION Preirus ultrasound machine. All of the screening program patients (191) with ultrasound-proven lesions underwent CB procedure and constituted 24% of this group.

## Statistical analysis

Mann-Whitney *U* test was used to assess differences between categories.  $\chi^2$  test was performed to compare the CB and VAB groups. Because of the number of cases, the comparison of BI-RADS 5 and 4a was not available, and comparison of the 4c group was done with  $\chi^2$  with Yates correction. All the statistical calculations were performed using Statistica software version 12.

Biopsy-proven predictive value (PPV3), which is the same as biopsy yield malignancy, also called positive biopsy rate, was calculated according to BI-RADS fifth edition.  $PPV3 = TP / (\text{number of biopsies})$ .

## Results

A total of 797 biopsies were analysed. A complete list of the types of histopathology is presented in Table 1.

BI-RADS 5 constituted 12% of cases (95), and BI-RADS 4 88% (698 cases); with BI-RADS 4 subdivisions there were 359 cases in 4a (45.3%), 215 in 4b (27%), and 124 in 4c (15.6%).

The number of diagnosed cancers was 276, which made up 34.8% of biopsies. In the BI-RADS 5 category, 93 cancers were detected from a total of 95 biopsies, and the

**Table 1.** Histopathological outcomes

Histopathology	No. of cases
Invasive ductal carcinoma	248
Fibroadenoma	190
Fibrocystic changes	168
Papilloma	33
Usual ductal hyperplasia	26
Invasive lobular carcinoma	17
Simple adenosis	15
Adenosis sclerosants	12
Columnar cell changes	12
Ductal carcinoma in situ	11
Hamartoma	7
Intramammary lymph node	6
Mastitis	6
Atypical ductal hyperplasia	5
Radial scar	5
Steatonecrosis	5
Atheroma	4
Duct ectasia	4
Lipoma	3
Atypical lobular hyperplasia	2
Flat epithelial atypia	2
Gynaecomastia	2
Haematoma	2
Abrikosow tumour	1
Apocrine metaplasia	1
Inflammatory granuloma	1
Lactating adenoma	1
Lymphoproliferative tissue	1
Metastatic lung cancer lymph node	1
Pseudoangiomatous stromal hyperplasia	1
Phyllodes tumour	1

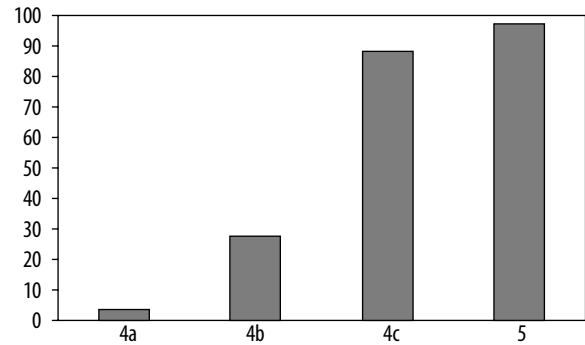
PPV3 was 97.9%. In the BI-RADS 4 category, PPV3 was 26.22% (183 cancers and 698 benign), in the division for 4a, 4b, and 4c subcategories the PPV3 values were 3.6%, 27.9%, and 88.7%, respectively (Table 2, Figure 1).

According to the type of biopsy, there were 248 cancers diagnosed by core biopsy (44.52% of CBs) and 28 by VAB (11.9% of VABs). In our study, the PPV3 differed between the CB and VAB groups (Table 3).

The most significant difference can be seen for the BI-RADS 4b category, and was of statistical importance, as shown by the  $\chi^2$  ( $p < 0.005$ ). The positive predictive value for BI-RADS 4b lesions was 38.1% vs. 13.5% in CB compared to VAB, respectively. According to  $\chi^2$  test with Yates correction, there is no statistically important difference

**Table 2.** Positive predictive value (PPV) according to BI-RADS category

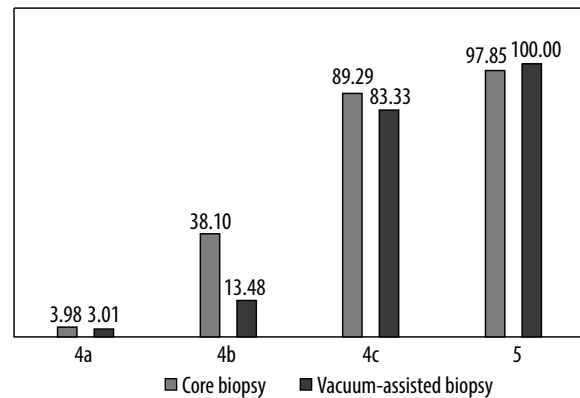
BI-RADS category	Total No.	Cancer No.	PPV 3
4a	359	13	3.63%
4b	215	60	27.91%
4c	124	110	88.71%
5	95	93	97.89%
Summary	793	276	34.80%



**Figure 1.** Percentage of diagnosed cancers according to BI-RADS category

**Table 3.** Positive predictive value according to BI-RADS category for core biopsy (CB) and vacuum-assisted biopsy (VAB)

BI-RADS	CB group	VAB group
4a	3.98%	3.01%
4b	38.10%	13.48%
4c	89.29%	83.33%
5	97.85%	100.00%



**Figure 2.** Percentage of diagnosed cancers

between the CB and VAB groups in the BI-RADS 4c category. The numbers of cases in category 4a and 5 were not suitable for statistical analysis. The PPV3 of the CB and VAB groups are presented in Figure 2. The Mann-Whitney *U* test proves the difference in the number of cancers between 4a, 4b, 4c, and 5 with  $p$  value  $< 0.005$ . The same test  $p$  value  $< 0.005$  was calculated for differences in the BI-RADS 4 subcategories. *U* Mann-Whitney  $p$  value  $< 0.005$  was obtained for the same differences separately calculated in the CB and VAB groups.

## Discussion

There were 31 types of histopathological outcomes (Table 1). Besides those expected, there were three, in our opinion, that required elaboration. The two of them are extramammary metastatic disease. Metastases are uncommon in breast and comprise no more than 2% of breast malignancy [17]. We diagnosed lymphoma and lung cancer as the most common in this group, of unknown origin [18]. Also of special concern is Abricossoff's tumour; a very rare benign neoplasm that can be found anywhere in the body, particularly the head and neck region, especially the tongue. Its location in the breast region is extremely rare – 6% of all granular cell carcinomas. On ultrasound, the mass appears solid, poorly defined, and with marked posterior shadowing [19]. BI-RADS categorisation in the studied group, especially in the case of BI-RADS 4c and 5, matches the literature data. BI-RADS 5 constitute 12.6% of cases, which is comparable with the 8% reported in "The Thai Journal of Surgery", especially as this was based only on nonpalpable mammographic lesions, and those limitations did not apply to our work [20]. ACR BI-RADS 4c was observed in 15.6% of cases and 4b in 27.1%, which is similar to the 12.6% and 31.8%, respectively, published by Elezaby *et al.* in "Radiology" in 2018 [9]. Compared with this assessment, there is clearly visible difference in the percentage of BI-RADS 4a. In our study, they constituted 45.3%, because in the quoted paper there were more than 55%. There might be several reasons for such a difference. The National Mammography Database is not limited to the type of biopsy guidance, so it consists of data from stereotactic biopsies as well. The number of suspicious calcifications seen on mammography is to be the BI-RADS 4b category, and in most cases those lesions are not visible in ultrasound examination, which excluded some of BI-RADS 4b lesions from our assessment and might be the reason for this difference. It should be mentioned that this is only a suspicion and should be verified.

Breast cancer was diagnosed in 34.8% of biopsies and in 26.2% in ACR BI-RADS 4 category. This is almost perfect pitch, as the BI-RADS 4 category reported range is from 20% to 27% and PPV for BI-RADS 5 and 4 in ultrasound-guided biopsy resume was 35% [9,20-25]. The category BI-RADS 4a is sometimes mentioned as a low cancer predictive factor, but because it is an indication for biopsy, we decided not to exclude it from our calculations for biopsy office, because it is still responsible for false positives and some true positives outcomes [26]. Mann-Whitney *U* test ( $p < 0.005$ ) revealed progressive increase of PPVs in the assessed categories and for BI-RADS 4a, b, and c subcategories alone ( $p < 0.005$ ), which is another confirmation of BI-RADS classification compliance [9,27]. The positive predictive value (PPV3) in BI-RADS 5 category was 97.89, which clearly matches the classification criteria and is adequate to previously reported, though there are some reports of lower PPV [20,28]. There were

only two cases of false positive BI-RADS 5 in our study, so we decided to look into the histopathology of those cases.

One proved to be atypical ductal hyperplasia and was further not upgraded after surgery.

The second false positive BI-RADS 5 lesion was an inflammatory atheroma.

Epidermoid cyst of the breast, when infected, can present all of the suspicious characteristics suggesting carcinoma, typical for BI-RADS 5 lesions, such as indistinct margins, echogenicity, and hyperechoic halo [29]. In the case of surrounding oedema, typical features like skin/subcutaneous tissue location or a thin neck extending to the skin, suggesting its benign origin, can be entirely masked. All of these features can lead to the need for histological evaluation.

Positive predictive value for category 4, as mentioned, in general was 26.2%. PPV3 of BI-RADS 4a category was quite low (3.63%), compared to 7.6% from the National Mammography Database.

On the other hand, the increase of PPV3 was observed in the remaining BI-RADS 4 categories.

For BI-RADS 4b PPV3 was calculated at 27.91% and was higher than that observed in the National Mammography Database, where it was 22%, but the difference is lower than in the case of the BI-RADS 4c category. In our study PPV3 for BI-RADS 4c lesion is 88.71% vs. 69.3% in the previously mentioned case. The secret may lie in the relatively low number of cases in our study in comparison to NMB, but it also might be due to the miscategorisation, because it is not clear if the shift would have been from category 4 or from 4a to BI-RADS 3.

There was, unaccepted for us, a difference in PPV3 for BI-RADS 4b in the CB and VAB group (Table 3). The PPV3 in the CB group is almost three times higher than in VAB (38.10% vs. 13.48%). There might be a number of reasons for this difference. Trying to find the reasons underlying this difference, we have to admit that there are different indications for those types of biopsies. In our facility masses bigger than 5 mm are referred to CB; on the other hand, cystic lesions with mass, classical BI-RADS 4b, are always referred to VAB.

The other difference is due to National Screening Program organisation. The program does not refund VAB; if there is need of one, the patient is referred to an oncological or surgical oncology outpatient clinic for the procedure. This results in high attendance of screening referrals in the CB group (191 from a total of 557). Because screening requirements (in Poland 5000 mammograms interpreted yearly) matches factor of interpretive performance, this might have an influence on the mentioned difference [11]. Due to the limitations of our hospital information system (HIS), we were unable to link the referring office with histopathological outcome, and, more importantly, after admission to the outpatient clinic the information about prior screening program attendance was lost. Although this is interesting, finding it should be under

further investigation, but in our case it must be done from screening examination office data, not from biopsy office data.

## Conclusions

The work has proven the stratification of cancer risk in BI-RADS biopsy-indicating criteria. It shows that among the biopt lesions the categories match those indicated in the

fifth edition of the ACR BI-RADS lexicon. An unacceptable difference in PPV3 between CB and VAB in category BI-RADS 4b was stated, which needs to be the subject of further investigation.

## Conflict of interest

The authors report no conflict of interest.

## References

- Breast imaging reporting and data system: BI-RADS atlas. 4<sup>th</sup> ed. American College of Radiology, Reston 2003.
- Costantini M, Belli P, Lombardi R, et al. Characterization of solid breast masses: use of the sonographic breast imaging reporting and data system lexicon. *J Ultrasound Med* 2006; 25: 649-659.
- D'Orsi CJ, Sickles EA, Mendelson EB, et al. ACR BI-RADS® Atlas. Breast Imaging Reporting and Data System. American College of Radiology, Reston 2013.
- Sang Yu N, Eun Young K, Boo-Kyung H, et al. Breast imaging reporting and data system category 3 lesions detected on whole-breast screening ultrasound. *J Breast Cancer* 2016; 19: 301-307.
- Sickles EA, D'Orsi CJ. ACR BI-RADS® Follow-up and Outcome Monitoring. In: ACR BI-RADS® Atlas. Breast Imaging Reporting and Data System. American College of Radiology, Reston 2013.
- D'Orsi CJ, Sickles EA. 2017 Breast Cancer Surveillance Consortium Reports on Interpretive Performance at Screening and Diagnostic Mammography: Welcome New Data, But Not as Benchmarks for Practice. *Radiology* 2017; 283: 7-9.
- Sardanelli F, Fallenberg EM, Clauser P, et al. Mammography: an update of the EUSOBI recommendations on information for women. *Insights Imaging* 2017; 8: 11-18.
- Jędrzyś S. Badania diagnostyczne. Zasady postępowania w zakresie wczesnego wykrywania raka piersi. *Medycyna Praktyczna* 2018. Available at: <https://www.mp.pl/onkologia/nowotwory/rak-piersi/diagnostyka/180174,wczesne-wykrywanie-raka-piersi> [In Polish].
- Elezaby M, Li G, Bhargavan-Chatfield M, et al. ACR BI-RADS Assessment Category 4 Subdivisions in Diagnostic Mammography: Utilization and Outcomes in the National Mammography Database. *Radiology* 2018; 287: 416-422.
- Ehsanbakhsh A, Toosi FS, Khorashadizadeh N, et al. Different BI-RADS Categories in Screening and Diagnostic Mammography. *Iran J Radiol* 2009; 6: 119-123.
- Miglioretti DL, Smith-Bindman R, Abraham L, et al. Radiologist characteristics associated with interpretive performance of diagnostic mammography. *J Natl Cancer* 2007; 99: 1854-1863.
- Ciatto S, Houssami N, Apruzzese A, et al. Reader variability in reporting breast imaging according to BI-RADS assessment categories (the Florence experience). *Breast* 2006; 15: 44-51.
- Chmielnik E, Łuczńska E. Biopsja gruboigłowa piersi – wytyczne diagnostyczne. *Pol J Pathol* 2009; 3 (Supplement 1): s13-s19.
- Perry N, Broeders M, de Wolf C, et al. European guidelines for quality assurance in breast cancer screening and diagnosis. 4<sup>th</sup> ed. Office for Official Publications of the European Communities, 2006.
- Bozzini A, Cassano E, Raciti D, et al. Analysis of Efficacy and Accuracy of 2 Vacuum-Assisted Breast Biopsy Devices: Mammotome and Elite. *Clin Breast Cancer* 2018; 18: e1277-e1282.
- Rozporządzenie Ministra Zdrowia z dnia 6 listopada 2013 r. Dz.U. 2913 poz. 1413.
- Conant EF, Brennecke CM. Breast imaging, case review series. Mosby Inc., 2006.
- Babu KS, Roberts F, Bryden F, et al. Metastases to breast from primary lung cancer. *J Thorac Oncol* 2009; 4: 540-542.
- Gavriliadis P, Michalopoulou I, Baliaka A, et al. Granular cell breast tumour mimicking infiltrating carcinoma. *BMJ Case Rep* 2013; pii: bcr2012008178.
- Wongmaneeerung P, Somwongprasert A, Watcharachan K, et al. Positive Predictive Value of BI-RADS 4 and 5 Mammography in Patients with Non-palpable Breast Mass. *Thai Journal of Surgery* 2015; 36: 116-119.
- Hooley R, Greenberg K, Stackhouse R, et al. Screening US in patients with mammographically dense breasts: initial experience with Connecticut Public. *Radiology* 2012; 256: 59-69.
- Poplack S, Carney P, Weiss J, et al. Screening mammography: costs and use of screening-related services. *Radiology* 2005; 234: 79-85.
- Mendez A, Cabanillas F, Echenique A, et al. Mammographic features and correlation with biopsy findings using 11-gauge stereotactic vacuum-assisted breast biopsy (SVABB). *Ann Oncol* 2004; 15: 450-454.
- Atasoy M, Tasali N, Çubuk R, et al. Vacuum-assisted stereotactic biopsy for isolated BI-RADS 4 microcalcifications: evaluation with histopathology and midterm follow-up results. *Diagn Interv Radiol* 2015; 21: 22-27.
- Hong A, Rosen E, Soo M, et al. BI-RADS for sonography: positive and negative predictive values of sonographic features. *AJR Am J Roentgenol* 2005; 184: 1260-1265.
- Rahimić-Čatić A, Bajrović J, Lozo J, et al. Sensitivity and specificity of magnetic resonance imaging in breast cancerous lesions prediction in the patients with BIRADS 4 and BIRADS 5 diagnoses. *Medical Journal* 2016; 22: 87-91.
- Ribas Do Nascimento JH, Duval da Silva V, Maciel AC. Accuracy of mammographic findings in breast cancer: correlation between BI-RADS classification and histological findings. *Radiolo Bras* 2010; 43: 91-96.
- Eun-Kyung K, Kyung H, Ki Keun O, et al. Clinical application of the BI-RADS final assessment to breast sonography in conjunction with mammography. *AJR Am J Roentgenol* 2008; 190: 1209-1215.
- Mendelson EB, Böhm-Vélez M, Berg WA, et al. ACR BI-RADS® ultrasound. In: ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System. American College of Radiology, Reston 2013.