

Can Cut-Off-Values for Tumor Size or Patient Age in Breast Ultrasound Reduce Unnecessary Biopsies or is it all About Bi-rads?– A Retrospective Analysis of 763 Biopsied T1-Sized Lesions



Authors

Laura Holzer-Fruehwald¹, Matthias Meissnitzer², Michael Weber³, Stephan Holzer⁴, Klaus Hergan⁵, Christian Weismann⁶

Affiliations

- 1 Department of Biomedical Imaging and Image-Guided Therapy, Medical University of Vienna, Vienna, Austria
- 2 Department of Radiology, University Hospital Salzburg PMU, Salzburg, Austria
- 3 Department of Biomedical Imaging and Image-Guided Therapy, Medical University of Vienna, Vienna, Austria
- 4 Department of Ophthalmology and Optometry, Medical University of Vienna, Vienna, Austria
- 5 Department of Radiology, University Hospital Salzburg PMU, Salzburg, Austria
- 6 Department of Radiology, University Hospital Salzburg PMU, Salzburg, Austria

Key words

breast, biopsy, 3D ultrasound

received 15.12.2016

revised 22.03.2017

accepted 13.04.2017

Bibliography

DOI <https://doi.org/10.1055/s-0043-110476>

Ultrasound Int Open 2017; 3: E94–E98

© Georg Thieme Verlag KG Stuttgart · New York

ISSN 2199-7152

Correspondence

Dr. Laura Holzer-Fruehwald, MD

Medizinische Universität Wien

Department of Radiology and Nuclear Medicine

Waehringerguertel 18-20/ertel 18-20

1090, Vienna

Austria

Tel: +43 1 40400 48180

laura.holzer-fruehwald@meduniwien.ac.at

ABSTRACT

Aims and Objectives To assess whether it is possible to establish a size cut-off-value for sonographically visible breast lesions in a screening situation, under which it is justifiable to obviate a biopsy and to evaluate the grayscale characteristics of the identified lesions.

Materials and Methods Images of sonographically visible and biopsied breast lesions of 684 patients were retrospectively reviewed and assessed for the following parameters: size, shape, margin, lesion boundary, vascularity, patient's age, side of breast, histological result, and initial BI-RADS category. Statistical analyses (t-test for independent variables, ROC analyses, binary logistic regression models, cross-tabulations, positive/negative predictive values) were performed using IBM SPSS (Version 21.0).

Results Of all 763 biopsied lesions, 223 (29.2%) showed a malignant histologic result, while 540 (70.8%) were benign. Although we did find a statistically significant correlation of malignancy and lesion size ($p = 0.031$), it was not possible to define a cut-off value, under which it would be justifiable to obviate a biopsy in terms of sensitivity and specificity (AUC: 0.558) at any age. Lesions showing the characteristics of a round or oval shape, a sharp delineation and no echogenic rim ($n = 112$) were benign with an NPV of 99.1%.

Conclusion It is not possible to define a cut-off value for size or age, under which a biopsy of a sonographically visible breast lesion can be obviated in the screening situation. The combination of the 3 grayscale characteristics, shape (round or oval), margin (circumscribed) and no echogenic-rim sign, showed an NPV of 99.1%. Therefore, it seems appropriate to classify such lesions as BI-RADS 2.

Introduction

There has been an ongoing debate about the application of ultrasound (US) in breast imaging for many years. The practice guidelines of the American College of Radiology (ACR) recommend the evaluation of palpable abnormalities or other breast symptoms, assessment of mammographic or MR imaging-detected abnormalities and evaluation of breast implants as the main indications for breast US [1]. Screening ultrasound (in addition to mammography) can be considered in high-risk patients for whom MRI may be inappropriate/contraindicated or in women with dense breast tissue [2–4]

US limitations include a limited field of view, high operator dependency and a sensitivity that is related to the size of the focal lesion [5, 6]. The inferior ability of US (compared to mammography) to depict microcalcifications reduces the sensitivity for ductal carcinoma in situ (DCIS) which is one of the factors preventing ultrasound from becoming a primary screening tool [4]. Multiple different settings of US which are adjusted during real-time scanning (choice of frequency, time gain compensation curve, focal zone depth, harmonic imaging, spatial compounding, and use of Doppler) can lead to very different impressions on static images. Thus, the main tasks of US remain the characterization and measurement of clinically and mammographically detected masses as well as guidance for interventions [7].

In 2003, the ACR introduced the first edition of the BI-RADS: Ultrasound lexicon, with the three most important mass characteristics being shape, margin and orientation [1, 8, 9]. The addition of US to mammography offers a clear cancer detection benefit in mammographically dense breasts (heterogeneously dense and extremely dense), leading to an incremental detection rate of small invasive, mostly node-negative cancers [2–4]. That makes ultrasound, in combination with mammography, the mainstay of breast imaging [10]. In January 2014, Austria started a population-based combined mammography/ultrasound breast cancer screening program. If the screening mammograms show dense parenchyma, additional screening ultrasound is immediately performed. The questions of our study are based on this additional ultrasound screening background, i.e. whether we can find a cut-off value for ultrasound lesion size within the T1 stage group (long axis diameter up to 2 cm) and/or patient age under which we can obviate a lesion biopsy. Or is it even possible to describe a combination of ultrasound lesion criteria which not only avoids a biopsy but may also prevent BI-RADS 3 short term follow-up?

Materials & Methods

Informed consent was waived by the local ethics committee as this was a retrospective study. We retrospectively reviewed the sonograms and histological reports of 684 women with 763 lesions who underwent US-guided biopsy at our department between 2003 and 2013. The asymptomatic patients visited our department either for the purpose of a screening exam or were referred for a second-look examination. 33 patients had 2 biopsies in the same breast at the same time, 8 patients had biopsies in both breasts at the same time and 4 patients had 3 biopsies at the same time either in one breast ($n = 1$) or in both breasts ($n = 3$). 19 patients had a biopsy twice in the same breast but at different time points and 10 patients had a biopsy in each breast at different time points.

For image acquisition and US-guided biopsy either an Aixplorer® with a linear-array transducer (SL 15-4 receiving transducer) (Supersonic imagine®, Aix-en-Provence, France) or a GE Voluson E8® with a SP 10-16D linear-array transducer (GE Healthcare Austria, Zipf, Austria) was used. Patients with sonographically detected BI-RADS 3 lesions (following the wish of the woman to rule out malignancy), BI-RADS 4 or 5 lesions and subsequent biopsy with a diameter between 1–20 mm (T1 stage lesion size) were included, when the lesion was documented showing its diameter in all three dimensions (in two orthogonal planes). As 3D images were not available in all patients, they were not used for evaluation. Patients with a history of breast cancer on either the ipsi- or contralateral breast and high-risk patients (BRCA-gene-positive tested patients) were not included in the study. All included patients were examined by the head of the breast imaging department (C.W., board-certified radiologist, 25 years of experience in breast radiology), who initially categorized the lesions according to BI-RADS and who performed all biopsies. The biopsies were performed using a 13-gauge coaxial cannula combined with a 14-gauge core needle and a Bard® Magnum® Reusable Core Biopsy System (Bard Biopsy Systems Tempe, USA), taking 3 to 5 cores (mean: 4 cores). Each probe was stored in an individual container ($n = 712$). In the case of very small lesions or complicated cysts, fine-needle aspirations using a 20-gauge needle were also performed instead of core biopsy ($n = 51$; size $2 \times 2 \times 3$ mm to $20 \times 13 \times 19$ mm). If a lesion was solid or it was not possible to aspirate, a core biopsy was subsequently performed.

The images of the biopsied lesions were reviewed (either on paper-print (2003–2008) or via a PACS system (2009–2013, AGFA IMPAX EE, Agfa HealthCare GmbH, Bonn, Germany) by one junior reader (LHF, after being trained by the senior reader for 4 months) and the following parameters were assessed: size (i.e., maximum lesion diameter), shape, margin, lesion boundary, vascularity, patient age (date of birth and date of biopsy), side of breast, histological biopsy result and – in case of surgery – also the final histopathological result. Surgery was performed in case of a malignant biopsy result and also in the case of at-risk lesions (see results section). Posterior features were not included because of their ultrasound technique dependency (frequency depending, depending on compound imaging which reduces the characteristics of posterior features compared without compounding). In the case of a malignant mass, the node status and the palpability was retrospectively assessed. Statistical analyses (t-test for independent variables, ROC analyses, binary logistic regression models, cross-tabulations, positive/negative predictive values) were performed using IBM SPSS (Version 21.0).

Results

The mean age of the women (diagnostic and screening) was 52.5 years (range: 20.9 to 90.5). The mean lesion diameter was 10 mm ($SD \pm 4.1$). Among 763 biopsied lesions, 223 (29.2%) proved malignant: the vast majority were invasive cancers ($n = 211$; 19.9% (42/211) node positive, 80.1% (169/211) node negative) with a median ultrasound diameter of 10.5 mm. The remaining 12 carcinomas were ductal carcinomas in situ (DCIS). Among the 223 malignant lesions, we found 58 (26%) G1, 143 (64.1%) G2 and 22 (9.9%)

► **Table 1** Rate of malignancy as a function of lesion size for 763 sonographically depicted breast masses 2 cm or smaller.

Size	Rate of malignancy n/total (%)
≤ 4 mm	3/41 (7.3%)
≤ 5 mm	10/98 (10.2%)
6–7 mm	52/149 (34.9%)
8–10 mm	63/214 (29.4%)
> 10 mm	98/302 (32.5%)
Overall	223/763 (29.2%)

G3 tumors. 540/763 (70.8%) masses were benign (among the benign cases we found: fibroadenomas (n = 118), complicated cysts (n = 88), papillomas (n = 25/surgery in 24 cases, 1 case showing early transition into a DCIS at final histopathology), ADHs (n = 4) and radial scars (n = 4)). The remaining 301 benign biopsies showed fibrous (or fibroadenomatous or fibrocystic) mastopathy (n = 236), foreign body granuloma (n = 1), liponecrotic cysts (n = 3), cholesterol granuloma (n = 3), lipoma (n = 1), chronic inflammation (n = 11), florid inflamed cyst (n = 2) or uncomplicated cysts, showing no dysplastic or atypical cells (n = 44) in the case of fine-needle aspiration.

The likelihood for malignancy increased with a greater lesion diameter (p = 0.031). In our series, we had no malignancies among masses ≤ 3 mm (n = 10), one of which was BI-RADS 3, six BI-RADS 4a, two BI-RADS 4b, and one BI-RADS 4 not further specified. 10 of 98 (10.2%) masses > 3 and ≤ 5 mm were malignant; masses measuring > 5 to ≤ 7 mm showed a malignancy rate of 34.9% (52 of 149), lesions > 7 to ≤ 10 mm were malignant in 29.4% (63 of 214) of cases and lesions > 10 mm showed a malignancy rate of 32.5% (98 of 302) ► **Table 1**.

The assessed BI-RADS descriptors proved to be reliable, showing malignancy in 1.6% of BI-RADS 4a masses (6 of 376) and 22.5% (23 of 102) and 60% (30 of 50) in the 4b and 4c categories, respectively. Of 166 BI-RADS 5 masses, 153 (92.2%) were malignant ► **Table 2**.

Binary logistic regression analysis showed a higher likelihood of malignancy the larger the lesion and the older the patient. The subsequent ROC analysis based on size only presented an AUC of 0.558 (see ► **Fig. 1**), showing that it is not justifiable to establish a cut-off value for size, below which it would be safe to avoid biopsy. The ROC analysis based on the combination of size and age also failed to demonstrate a reliable cut-off-value as the AUC was 0.751.

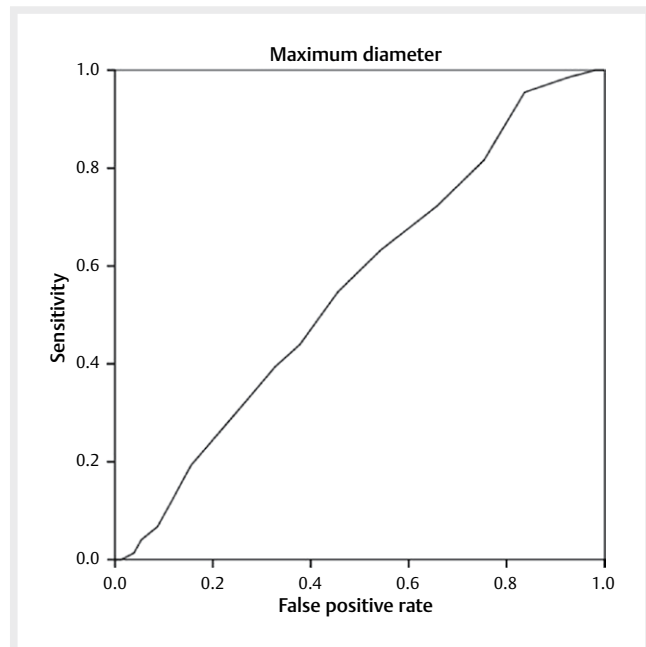
Regarding T1 lesions, those fulfilling the three criteria of having 1) an oval/round shape (NPV: 94.1%), 2) a circumscribed margin (NPV: 97.6%) and 3) no echogenic rim (NPV 81.1%), only 1 of 112 lesions (0.9%) proved malignant (a 10-mm grade 3 invasive ductal carcinoma in a 29-year-old woman with a palpable lump) making the sum of those criteria highly predictive of benignity with an overall NPV of 99.1% (0.951 to 0.998 at a 95% CI) – see ► **Table 3**, and a sensitivity of 99.6%.

The BI-RADS assessments of these 112 masses were: BI-RADS 3 (n = 11), BI-RADS 4 (n = 6), BI-RADS 4a (n = 83) and BI-RADS 4b (n = 12).

Among the benign lesions, we had 48 cases in which the biopsy result, although benign, still lead to surgery (because of a higher-

► **Table 2** Rate of malignancy as a function of BI-RADS assessment for 763 sonographically depicted breast masses 2 cm or smaller.

BI-RADS	Rate of malignancy n/total (%)
3	0/28 (0%)
4	11/41 (26.8%)
4a	6/376 (1.6%)
4b	23/102 (22.5%)
4c	30/50 (60%)
5	153/166 (92.2%)



► **Fig. 1** ROC curve based on diameter for malignant/benign histology: the larger the lesion, the more likely a malignant result. However, with an AUC of 0.558, it is not possible to demonstrate a reliable cut-off value.

► **Table 3** Sonomorphologic lesion features and their risk of malignancy.

Sonomorphologic lesion features	Negative predictive value (NPV)
Oval/round shape	94.1%
Circumscribed margin	97.6%
No echogenic rim	81.1%
Overall	99.1%

risk lesion like a papilloma (n = 35), atypical cells (n = 7) or radial scars (n = 5)). In one patient in whom surgery was recommended despite a benign biopsy result because of the suspicious sonomorphological appearance, histopathology of the surgically removed

specimen revealed cancer (invasive-papillary carcinoma G1 with intraductal components).

Discussion

In this study we wanted to analyze the influence of lesion size alone and in combination with patient age to find a “rule of thumb” for the combined mammography-ultrasound screening setting. We did not try to establish a new kind of prediction model like other authors have designed [11]. Instead we assessed whether it was possible to recommend a cut-off value for size (and/or age), under which it is reasonable to obviate a needle biopsy. Although we found statistically significant correlations between malignancy rate and tumor size and patient age respectively in the binary logistic regression analysis, it was not possible to establish a cut-off-value for those criteria while retaining an acceptable sensitivity and specificity. Paulinelli et al. did not find any tumor size that could be significant in the multivariate analyses, but did find an increasing risk of malignancy with increasing age (ROC giving a sensitivity of 84.5%, a specificity of 64.5% and an overall accuracy of 74.5% for the age of 40) [11].

We found that the combination of oval/round shape and circumscribed margin without echogenic rim (= circumscribed margin in the 5th edition of the ACR BIRADS® Atlas) [8] was the most reliable to predict a benign outcome with a negative predictive value (NPV) of 99.1%. In a study by Stavros, the NPV (99.5%) and sensitivity (98.4%) were even higher, but they evaluated 20 different sonographic features, in contrast to our study in which we concentrated on only 3 characteristics. Hong et al. [12] showed an NPV of 84% for lesions of oval shape, an NPV of 90% for circumscribed margins and 87% for parallel orientation. Baez et al. [13] described a round or oval shape, circumscribed margins, the lack of architectural distortion and no edge refraction/spiculation as the features with the highest negative predictive value (95.1, 98.2, 96.2 and 95.3% respectively). We draw the conclusion that it is reasonable to evaluate these 3 grayscale criteria above everything else, independent of family history, lesion size and age of the patient.

This conclusion leads to the suggestion that lesions fulfilling those criteria should not be needle biopsied and therefore not be rated as BI-RADS 4a, but as BI-RADS 3 and possibly as BI-RADS 2 in a screening population with normal risk in order to reduce the number of BI-RADS 3 short-term follow-ups. Our only case showing a malignant histology despite having an oval appearance with smooth margins and a lack of an echogenic rim was a 10 mm large, palpable lump in a 29-year-old patient (node positive, invasive ductal carcinoma, G3). Concerning the decision if it is more appropriate to classify lesions as BI-RADS 2 or 3, Barr et al. suggest follow-up of BI-RADS 3 lesions detected at screening US with no suspicious features at 1 year [14], and not – as usually required by BI-RADS 3 category – after 6 months. They thus propose a kind of procedure “in between” the two categories as they do not see a benefit in sonographic follow-up at 6 months (in their study 0.1% of the cancers had suspicious changes at 6 months, and the malignancy rate of BIRADS 3 lesions was 0.8%). Another study by Graf and colleagues showed an NPV of 99.8% for BI-RADS 3 classified lesions and 99.3% of the evaluated lesions remained stable during 2 years of follow-up. The one malignant lesion they had in their

series became palpable and was therefore clinically suspicious after 4 months (before the follow-up exam was due) [15]. Kim et al. assessed the outcome of asymptomatic category 2–4 BIRADS lesions by follow-up (< 12 months) or biopsy and found an NPV for the BI-RADS 2 category of 100% and an NPV of 99.2% for category 3 lesions [16]. Those 2 studies showing an NPV of more than 99% for lesions classified as BI-RADS 3 also give cause to debate if it is appropriate to classify such lesions directly as BI-RADS 2, but with diagnostic rather than routine screening follow-up.

In our series, we found a relatively low malignancy rate in BI-RADS category 4a (1.6%). This may reflect a bias towards biopsy rather than follow-up at our institution. In a study evaluating the interobserver variability for sonographic descriptors, Lazarus et al. found substantial agreement for assessment of lesion orientation ($\kappa = 0.61$), substantial agreement for evaluation of lesion shape ($\kappa = 0.66$) (“irregular” and “oval” shape showing the highest kappa-values of $\kappa = 0.70$ and 0.71 , respectively) and also substantial agreement for evaluation of the lesion boundary being abrupt or showing an echogenic halo/rim ($\kappa = 0.69$) [17]. Evaluation of lesion margin showed only fair agreement of $\kappa = 0.40$ overall, but describing circumscribed lesions yielded very high kappa-values ($\kappa = 0.71$). In this study the positive predictive values of the different BI-RADS categories were also calculated, showing similar results as in our study (6% for BI-RADS 4a, 15% for BI-RADS 4b, 53% for BI-RADS 4c and 91% for BI-RADS 5). Another study also found the highest inter- and intraobserver variability for mass shape ($\kappa = 0.8$ and 0.79 , respectively) [18].

While viewing the images of the included study objects, we also looked at the vascularity, but decided not to statistically evaluate this feature, as it is too user-dependent (considering that the type of Doppler mode [color versus power Doppler] as well as the applied pressure of the transducer and the adjusted sensitivity of the machine play an important role and the lack of measured resistance or pulsatility index in our study population did not allow a standardized way of comparing vascularity). Moreover, we would have expected diverse results, as intratumoral blood flow might be decreased based on a higher intratumoral pressure in malignant lesions due to the irregular vascular pattern and the occurrence of pathological vessels, while the loss of intratumoral tissue elasticity may lead to an increase in blood flow resistance [10]. A review by Hooley et al. summarizes the diverse results on Doppler imaging studies [19].

Limitations

We demonstrated a correlation between lesion size and risk for malignancy of a lesion. However, this may be biased by the fact that smaller cancers may more easily be overlooked compared to benign smaller lesions such as lesions with cystic components or biased by the fact that it is more difficult to correctly needle biopsy a lesion measuring 2 or 3 mm compared to larger lesions.

We did not assess the palpability of all lesions in retrospect (only for the malignant lesions) or if a lesion was new or increasing in size compared to the previous examination.

Clinical relevance

We could not define a cut-off value for size or age in the T1 lesion size group, below which needle biopsy of a sonographically visible breast lesion can be obviated (under screening conditions without

palpable masses). The combination of the grayscale characteristics, shape (oval/round) and circumscribed margin (well defined and no echogenic-rim sign), showed an NPV of 99.1 % (0.951 to 0.998 at a 95 % CI) so it seems appropriate to classify such lesions as BI-RADS 2. This may help to reduce the number of unnecessary follow-up exams.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. ACR BI-RADS® atlas. ACR BI-RADS® atlas, breast imaging reporting and data system: Available from. <http://www.acr.org/Quality-Safety/Resources/BIRADS/About-BIRADS/How-to-Cite-BIRADS> Accessed 19 February 2015
- [2] Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: An analysis of 27,825 patient evaluations. *Radiology* 2002; 225: 165–175
- [3] Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Böhm-Vélez M et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. *JAMA* 2008; 299: 2151–2163
- [4] Lee CH, Dershaw DD, Kopans D, Evans P, Monsees B, Monticciolo D et al. Breast cancer screening with imaging: Recommendations from the society of breast imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. *J Am Coll Radiol* 2010; 7: 18–27
- [5] Berg WA, Blume JD, Cormack JB, Mendelson EB. Operator dependence of physician-performed whole-breast US: Lesion detection and characterization. *Radiology* 2006; 241: 355–365
- [6] Cilotti A, Bagnolesi P, Moretti M, Gibilisco G, Bulleri A, Macaluso AM, Bartolozzi C. Comparison of the diagnostic performance of high-frequency ultrasound as a first- or second-line diagnostic tool in non-palpable lesions of the breast. *Eur Radiol* 1997; 7: 1240–1244
- [7] Benson SR, Blue J, Judd K, Harman JE. Ultrasound is now better than mammography for the detection of invasive breast cancer. *Am J Surg* 2004; 188: 381–385
- [8] Mercado CL. BI-RADS update. *Radiol Clin North Am* 2014; 52: 481–487
- [9] Burnside ES, Sickles EA, Bassett LW, Rubin DL, Lee CH, Ikeda DM et al. The ACR BI-RADS experience: Learning from history. *J Am Coll Radiol* 2009; 6: 851–860
- [10] Weismann CF. Breast ultrasound: New frontiers in imaging? *Ultrasound Obstet Gynecol* 2000; 15: 279–281
- [11] Paulinelli RR, Freitas-Junior R, de Lucena CÊ, Moreira MA, de Moraes VA, Bernardes-Júnior JR et al. Sonobreast: Predicting individualized probabilities of malignancy in solid breast masses with echographic expression. *Breast J* 2011; 17: 152–159
- [12] Hong AS, Rosen EL, Soo MS, Baker JA. BI-RADS for sonography: Positive and negative predictive values of sonographic features. *AJR Am J Roentgenol* 2005; 184: 1260–1265
- [13] Baez E, Strathmann K, Vetter M, Madjar H, Hackelöer BJ. Likelihood of malignancy in breast lesions characterised by ultrasound with a combined diagnostic score. *Ultrasound Med Biol* 2005; 31: 179–184
- [14] Barr RG, Zhang Z, Cormack JB, Mendelson EB, Berg WA. Probably benign lesions at screening breast US in a population with elevated risk: Prevalence and rate of malignancy in the ACRIN 6666 trial. *Radiology* 2013; 269: 701–712
- [15] Graf O, Helbich TH, Hopf G, Graf C, Sickles EA. Probably benign breast masses at US: Is follow-up an acceptable alternative to biopsy? *Radiology* 2007; 244: 87–93
- [16] Kim SJ, Chang JM, Cho N, Chung SY, Han W, Moon WK. Outcome of breast lesions detected at screening ultrasonography. *Eur J Radiol* 2012; 81: 3229–3233
- [17] Lazarus E, Mainiero MB, Schepps B, Koelliker SL, Livingston LS. BI-RADS lexicon for US and mammography: Interobserver variability and positive predictive value. *Radiology* 2006; 239: 385–391
- [18] Baker J, Kornguth PJ, Soo MS, Walsh R, Mengoni P. Sonography of solid breast lesions: Observer variability of lesion description and assessment. *AJR Am J Roentgenol* 1999; 172: 1621–1625
- [19] Hooley RJ, Scoutt LM, Philpotts LE. Breast ultrasonography: State of the art. *Radiology* 2013; 268: 642–659