Interobserver Variability and Distribution of Mammographic Breast Density in Patients with Invasive Breast Carcinoma using the Fifth Edition of BI-RADS Guidelines

Ab Rahman SS*[a], Mohd Kamal LH[b], Ab Rahman FH[c], Ghazali R[d], Saidil E[e], Mohamad Nor N[f]

[a]Radiology Unit, Faculty of Medicine and Health Sciences, Universiti Sains Islam Malaysia

[b]Faculty of Medicine and Health Sciences, Universiti Sains Islam Malaysia

[c]Surgery Unit, Faculty of Medicine and Health Sciences, Universiti Sains Islam Malaysia

[d]Department of Radiology, Hospital Ampang

[e]Community Health Unit, Faculty of Medicine and Health Sciences, Universiti Sains Islam Malaysia

ABSTRACT

INTRODUCTION: Breast density is associated with an increased risk of developing breast cancer. The present study aims to determine the distribution and interobserver variability of mammographic breast density in patients with invasive breast carcinoma, using the fifth edition of BI-RADS guidelines. It is part of a larger study to ascertain the association between mammographic breast density and breast cancer characteristics. **MATERIALS AND METHODS**: Two radiologists independently assessed 122 mammograms of patients with histologically confirmed invasive breast carcinoma and assigned the breast density to categories A-D based on the fifth edition of BI-RADS guidelines. The interobserver variability was calculated using the weighted kappa coefficient and the level of agreement was determined using the Landis and Koch guidelines. **RESULTS**: In this study, 55.7% of patients with invasive breast carcinoma were assigned to category B, followed by category C with 36.1%. Only 4.1% of patients were assigned to categories A and D respectively. There was substantial agreement between the two readers' judgement, k=0.610 (95% CI, 0.523-0.697), $p < 0.001$ for specific BI-RADS categories. **CONCLUSION**: Among patients with invasive breast carcinoma, there were more patients with non-dense breasts than dense breasts. Overall, there is a substantial interobserver agreement when radiologists used the fifth edition of the BI-RADS guideline, which is in line with results found in the literature. This suggests that the BI-RADS density classification is an acceptable method and can be reliably used in clinical practice.

INTRODUCTION

Breast cancer is the most commonly diagnosed cancer and the leading cause of cancer-related death among women worldwide.[1] In Malaysia, a total of 3,766 new breast cancer cases were reported in 2011, with an overall Age-Standardised Incidence Rate (ASR) of 31.1 per 100,000 population.[2] Breast density is associated with an increased risk of developing breast cancer.[3,4] Increased mammographic breast density reflects a higher percentage of dense fibroglandular tissue compared to radiolucent fatty tissue. Several studies have shown that breast density is an independent risk factor for breast cancer, and is associated with up to six-fold increased risk.[5,6,7] Furthermore, the appearance of dense breast tissue makes it harder to detect breast cancer due to the masking effect, thus limiting the sensitivity of mammography.[7,8]

Several methods, either qualitative or quantitative, have been proposed to measure breast density, including classification systems detailed by Wolfe[9], Boyd et al.[10], Tabar[11], and commercially available automated volumetric breast density software. Quantitative systems are more objective, reliable, and reproducible, compared with qualitative methods which are more

subjective and reliant on individual judgement.[12,24] However, in many clinical settings, predominantly in developing countries, there is a lack of quantitative systems, which calls for improvement in the reproducibility of qualitative methods. The American College of Radiology (ACR) Breast Imaging Reporting and Data System (BI-RADS) lexicon was developed to standardise breast imaging reports, enhance communication with referring clinicians, and deliver a quality assurance tool,[13] and is one of the widely used methods for assessing mammographic breast density.

Published in 2013, the fifth edition of BI-RADS classified the mammographic density into categories A, B, C, and D (Figure 1). A: the breasts are almost entirely fatty; B: there are scattered areas of fibroglandular density; C: the breasts are heterogeneously dense, which may obscure small masses; D: the breasts are extremely dense, which lowers the sensitivity of mammography.[14] The present study aims to determine the distribution and interobserver variability of the mammographic density in patients with invasive breast carcinoma, using the fifth edition of BI-RADS guidelines.
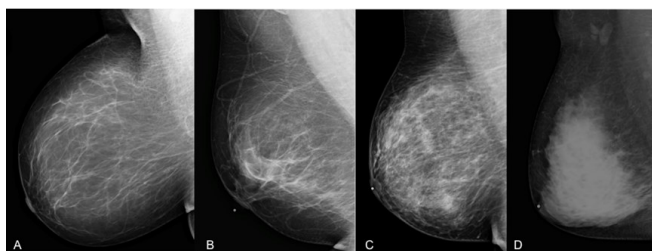

Figure 1: MLO views mammography of four different patients demonstrating the four categories of mammographic density based on the fifth edition of BI-RADS. A: the breast is almost entirely fat; B: there are scattered areas of fibroglandular density; C: the breast is heterogeneously dense; D: the breast is extremely dense.

MATERIALS AND METHODS

Study design

This study is part of a larger study to determine the association between mammographic breast density and breast cancer characteristics. The Medical Research and Ethics Committee of Malaysia approved the study, and informed consent was waived as the study was of a retrospective nature. Within the study period of 2014 to 2017, 168 histologically confirmed invasive breast carcinoma cases were collected from the hospital database. Out of these, 122 cases had digital mammogram images available. The women's ages

ranged from 32 – 82 years of age (mean: 47 years).

The breast density was assessed from the mediolateral oblique (MLO) and craniocaudal (CC) images of the cancer-free breast, and assigned the density into categories A – D. Two general radiologists who were blinded to the original density reports independently evaluated the breast density of every case, and all disagreements were resolved with consensus. Both radiologists had more than 5 years of clinical experience in reporting mammograms, and specifically 2 years of experience using the fifth edition of BI-RADS guidelines. A consensus meeting was held at the beginning of this study between both radiologists to discuss the ACR guidelines and to ensure all uncertainties were addressed. For analysis purposes, BI-RADS A and B categories were regarded as "non-dense breast" while BI-RADS C and D categories were regarded as "dense breast".

Statistical analysis

The interobserver variability between the two radiologists was calculated using Fleiss-Cohen weighted kappa coefficient (k) and its 95% confidence interval (95% CI) since breast density was rated on an ordinal scale. The cells that were further from the agreement were weighted lower than those closer in agreement.[15]

Based on the k values, the Landis and Koch guidelines were applied to determine the level of agreement. A kappa value of 1.0 was regarded as perfect agreement, while a kappa value of 0 was regarded as no agreement. The k value was further classified as follows: 0.01 was considered poor agreement; 0.02–0.20 was considered slight agreement; 0.21–0.40 was considered fair agreement; 0.41–0.60 was considered moderate agreement; 0.61–0.80 was considered substantial agreement; and 0.81–0.99 was considered almost perfect agreement.[16] All analyses were carried out using statistical software (IBM SPSS, version 24.0).

RESULTS

Mammographic density distribution

A total of 122 mammograms were assessed, and the distribution of breast density based on the final

consensus is as follows: 68 patients (55.7%) were assigned to category B, followed by category C with 44 patients (36.1%). Only 5 patients (4.1%) were assigned to categories A and D respectively. Seventy-three patients (59.8%) were assigned as non-dense breasts while 49 patients (40.2%) were assigned as dense breasts. Figure 2 shows the frequency distribution of the cases classified as BI-RADS categories A–D by each reader.
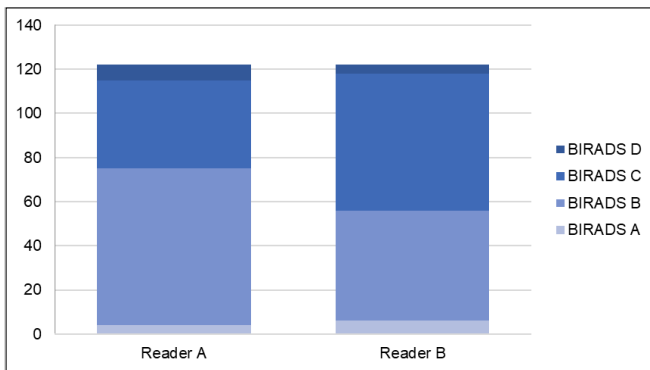

Figure 2: Frequency of distribution of cases categorised as BI-RADS A – D by the two readers.

Table 1 showed the agreement between the two readers according to BI-RADS categories and types of breast density. For specific BI-RADS categories, the agreement rate among readers was 100.0% for category A, 69.0% for category B, 95.0% for category C, and 42.9% for category D. For non-dense and dense categories, the agreement rate among readers was 98.2%, and 69.7% respectively. For specific BI-RADS categories (k = 0.610, $p < 0.001$) and types of breast density (k= 0.610, $p < 0.001$) , there were a significant substantial agreement between the two readers respectively.

Interobserver agreement

Variables		Number of agreement	% of agreement	Kappa value	p-value
				Reader B	
BI-RADS categories					
BI-RADS A	Reader A	4	100.0		
BI-RADS B		49	69.0	0.612	<0.001
BI-RADS C		38	95.0		(CI: 0.523-0.697)
BI-RADS D		3	42.9		
Types of breast density			Reader B		
Nondense	Reader A	55	98.2	0.662	<0.001
Dense		46	69.7		(CI: 0.537-0.787)

Table 1: Agreement between the two readers according to BI-RADS categories (A-D) and types of breast density.

DISCUSSION

Breast density is an important component in mammogram evaluation, so it is vital to have a breast density assessment system that is not only accurate and reliable but also consistent and reproducible. The visual estimation of breast density may be affected by individual perceptual differences, as well as technical and positional factors as it is not a precise science.[7] From our study, we calculated that the interobserver agreement for specific BI-RADS categories is substantial (k=0.610, 95% CI, 0.523 – 0.697). This is comparable with other studies that evaluated the interobserver agreement using the fifth edition BI-RADS guideline, which reported k values ranging from 0.57 to 0.79.[5,7,12] The summary of these studies is detailed in Table I. The studies differ in the number of radiologists and also the number of mammograms evaluated.

Ekpo et al. studied the interobserver variability between 5 radiologists who read a test set of 1000 mammograms and reported substantial interobserver agreement in the measurement of the mammographic breast density (k=0.79; 95% CI, 0.78-0.83) on a four-category scale (categories A-D).[12]

Several studies have compared the interobserver agreement between the fifth and fourth editions of BI-RADS to determine its consistency and reliability in clinical practice. In the study by Afsaneh et al., 3 radiologists reviewed a total of 72 mammograms four times; twice using the respective BI-RADS edition with each review separated by a one-month gap and its review order changed. The interobserver agreement for the fourth and fifth BI-RADS editions was 0.623 (95% CI, 0.517–0.729) and 0.702 (95% CI, 0.589-0.815) respectively, suggesting substantial agreement.[5] Irshad et al. also compared the interobserver agreement using the fourth and fifth edition of BI-RADS and reported a drop in the kappa values from substantial (k=0.65, 95% CI, 0.61-0.69) using the fourth-edition BI-RADS to moderate (k=0.57, 95% CI, 0.53-0.61) using the fifth edition and the difference in these values was statistically significant (p=0.006).[7]

Few other studies compared the interobserver variability between readers who interpret breast density subjectively

Table I: Summary of studies researching the interobserver variability using the fifth edition of BI-RADS.

No.	Title	Country	Number of mammograms reviewed	Number of readers	Classification	Interreader agreement
1	Assessment of Interradiologist Agreement Regarding Mammographic Breast Density Classification Using the Fifth Edition of the BI-RADS Atlas. (Ekpo et al, 2016)	Nigeria, Australia	1000	5	Fifth Edition BI-RADS	k= 0.79 (95% CI, 0.78-0.83) Substantial agreement
2	Comparison of inter-and intra-observer variability of breast density assessments using the fourth and fifth editions of Breast Imaging Reporting and Data System. (Alikhassi et al, 2018)	Iran	72	3	Fourth Edition BI-RADS Fifth Edition BI-RADS	k= 0.6023 (95% CI, 0.517 – 0.702) Substantial agreement k= 0.702 (95% CI 0.589 – 0.815) Substantial agreement with increased k value (p=32)
3	Effects of Changes in BI-RADS Density Assessment Guidelines (Fourth Versus Fifth Edition) on Breast Density Assessment: Intra- and Interreader Agreements and Density Distribution. (Irshad et al, 2016)	United States	104	5	Fourth edition BI-RADS Fifth Edition BI-RADS	k= 0.65 (95% CI, 0.61-0.69) Substantial agreement k= 0.57 (95% CI, 0.53-0.61) Moderate agreement with decreased k value. (p=0.06)
4	Comparison of variability in breast density assessment by BIRADS category according to the level of experience. (Eom et al, 2018)	Korea	1000	6	Fifth Edition BI-RADS	k=0.67 (95% CI, 0.63-.070) Substantial agreement
5	The inter-observer variability of breast density scoring between mammography technologists and breast radiologists and its effect on the rate of adjuvant ultrasound. (Mazor et al, 2016)	Israel	503	17	Fifth Edition BI-RADs	k=0.69 (95% CI, 0.59-0.78) Substantial agreement

using the fifth edition of BI-RADS and comparing it with objective, automated volumetric assessments of breast density. Eom et al. studied the difference in breast density assessment according to BI-RADS categories between readers with varying levels of experience and compared it with the findings of automated volumetric assessments. They found that the agreement between the visual and volumetric assessment for breast imaging experts and the general radiologist was substantial (k= 0.71-0.77).[22] Ekpo et al. studied breast density assessment using quantitative software and determined its agreement with radiologists' assessment using the fifth edition BI-RADS. Interobserver agreement ranged from fair (0.38, 95%CI, 0.30-0.46) to substantial (0.68, 95%CI, 0.61-0.75) on a four-category scale using the fifth edition of BI-RADS.[23]

In our study population, 40.2% of the patients with invasive breast carcinoma were categorised as having dense breasts, i.e. categories C and D. Jiang et al reported an almost similar distribution in their cohort, with 47.2% of patients reported to have dense breasts.[19] However, their classification was based on the fourth edition of BI-RADS. Our study found that 55.7% of the patients with invasive breast cancer were of category B density. A study by Gill et al used computed-aided software to quantify the percentage of breast density and found that the mean breast density in their invasive breast cancer population was 36.5%.[20] Ko et al also used computed-aided software to classify the breast density in their study population and reported a mean breast density of 21.1% among the invasive breast carcinoma patients.[21]

Increased mammographic breast density is associated with a higher occurrence of interval cancers.[24] Interval cancers are cancers that are detected between regular screening examinations and considered as false negatives, therefore lowering the sensitivity of mammography.[24] Mammographic breast density is also recognised as an independent risk factor for breast cancer, with a relative risk of 4- to 6-fold for dense breasts as compared to non-dense breasts.[24,25] Different techniques of measuring mammographic breast density produce varying degrees of association with breast cancer. Although there is significant overlap in the risk associations between these techniques, volumetric methods have been shown to give the strongest association.[24]

A limitation of this study is the relatively small number of radiologists. The readers were general radiologists who did not receive formal training in the fifth edition of BI-RADS prior to the initiation of the study but were working in the same clinical setting. Furthermore, fewer patients were belonging to BI-RADS categories A and D, resulting in an uneven dataset. The present study required radiologists to solely concentrate on breast density, which may have been a contributing factor to the high interobserver agreement. However, in actual clinical practice, breast density is normally not the main focus of the report. Therefore, the generalizability of the results may be limited.

Contrarily, the strength of the study is that breast density evaluation was conducted by radiologists who are currently using the BI-RADS guidelines in everyday clinical work. The number of mammograms reviewed was adequate to calculate the statistical accuracy of interobserver agreement.

CONCLUSION

This study demonstrated that among patients with invasive breast carcinoma, there were more patients with non-dense breasts compared to dense breasts. There is a substantial interobserver agreement when radiologists used the fifth edition of the BI-RADS guideline, which is conforming to results reported in the literature. This suggests that the classification is an acceptable method and can be reliably applied in clinical practice.

REFERENCES

1. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. CA Cancer J Clin 2011; 61:134.
2. Azizah AM, Nor Saleha IT, Noor Hashimah A, Asmah ZA, Mastulu W. Malaysian National Cancer Registry Report 2007-2011: National Cancer Institute, Ministry of Health Malaysia, 2015 June.
3. Erin JA, Diana SMB, Emily W, Peggy LP. Association between mammographic breast density and breast cancer tumor characteristics. Cancer Epidemiol Biomarkers Prev 2005; 14: 662-8.
4. Freer PE. Mammographic breast density: impact on breast cancer risk and implications for screening. RadioGraphics 2015; 35: 302-315.
5. Afsaneh A, Hamed EG, Masoud B. Comparison of inter- and intra-observer variability of breast density assessments using the fourth and fifth editions of Breast Imaging Reporting and Data System. Eur J Radiol Open 2018; 5: 67-72.
6. Tesic V, Kolaric B, Znaor A, Kuna SK, Brkljacic B. Mammographic density and estimation of breast cancer risk in intermediate risk population. Breast J 2013; 19: 71-8.
7. Abid I, Rebecca L, Susan A, et al. Effects of changes in BI-RADS density assessment guidelines (fourth versus fifth edition) on breast density assessment: intra- and interreader agreements and density distribution. Am J Roentgenol 2016; 207: 1366-71.
8. Pisano ED, Hendrick RE, Yaffe MJ, et al. Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. Radiology 2008; 246: 376-83.
9. Wolfe JN. Breast patterns as an index of risk for

developing breast cancer. Am J Roentgenol 1976; 126: 1130-7.

10. Boyd NF, Byng JW, Jong RA, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. J. Natl. Cancer Inst 1995; 87: 670-5.

11. Gram IT, Funkhouser E, Tabar L. The Tabar classification of mammographic parenchymal patterns. Eur. J. Radiol 1997; 24: 131-6.

12. Ernest UE, Ujong PU, Claudia MT, Mark FM. Assessment of interradiologist agreement regarding mammographic breast density classification using the fifth edition of the BI-RADS atlas. Am J Roentgenol 2016; 206: 1119–23.

13. Rao AA, Feneis J, Lalonde C, Ojeda-Fournier H. A pictorial review of changes in the BI-RADS Fifth Edition. RadioGraphics 2016; 36: 623-39.

14. Sickles EA, D'Orsi CJ, Bassett LW, Appleton CM, Berg WA, Burnside ES. ACR BI-RADS Atlas, Breast Imaging Reporting and Data System. Reston, VA: American College of Radiology; 2013.

15. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educ. Psychol. Meas 1973; 33: 613-9.

16. Landis JR, Koh GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33: 159-174.

17. Alowami S, Troup S, Al-Haddad S, Kirkpatrick I, Watson PH. Mammographic density is related to stroma and stromal proteoglycan expression. Breast Cancer Res 2003; 5: R129 – 35.

18. Byrne C, Colditz GA, Willett WC, Speizer FE, Pollak M, Hankinson SE. Plasma insulin-like growth factor (IGF) I, IGF-binding protein 3, and mammographic density. Cancer Res 2000; 60: 3744 -8.

19. Liyu J, Tingting MA, Meena SM, Xiaoli K, Xiaoyan L, Bruce G, et al. Mammographic features are associated with clinicopathological characteristics in invasive breast cancer. Anticancer Res 2011; 31: 2327-34.

20. Gill JK, Maskarinec G, Pagano I, Kolonel LN. The association of mammographic density with ductal carcinoma in situ of the breast: the Multiethnic Cohort. Breast Cancer Res 2006; 8: R30.

21. Ko H, Shin J, Lee JE, et al. Comparison of the association of mammographic density and clinical factors with ductal carcinoma in situ versus invasive breast cancer in Korean women. BMC Cancer 2017; 17: 821.

22. Eom HJ, Cha JH, Kang JW, Choi WJ, Kim HJ, Go E. Comparison of variability in breast density assessment by BI-RADS category according to the level of experience. Acta Radiologica 2018; 59: 527 -32.

23. Ekpo EU, Mello-Thoms C, Rickard M, Brennan PC, McEntee MF. Breast density (BD) assessment with digital breast tomosynthesis (DBT): agreement between Quantra™ and 5th edition BI-RADS. The Breast 2016; 30: 185-190.

24. Destounis S, Arieno A, Morgan R, Roberts C, Chan A. Qualitative versus quantitative mammographic breast density assessment: applications for the US and abroad. Diagnostics 2017; 7: 30.

25. Nazari SS, Mukherjee P. An overview of mammographic density and its association with breast cancer. Breast Cancer 2018; 25: 259–267.