European Journal of Therapeutics

pISSN: 2564-7784
eISSN: 2564-7040

Eur J Ther. 2024
https://doi.org/10.58600/eurjther2002

Original Research

Evaluation of Automated Mammographic Density Classification in Tomosynthesis: Comparison with Radiologists

Hüseyin Alper Kızıloğlu [1], Murat Beyhan [1], Erkan Gökçe [1], Yaşar Birişik [1], Muhammet Furkan Battal [1], Muhammed Erkam Çeker [1], Osman Demir [2]

[1] Department of Radiology, Tokat Gaziosmanpaşa University Faculty of Medicine, Tokat, Türkiye
[2] Department of Biostatistics, Tokat Gaziosmanpaşa University Faculty of Medicine, Tokat, Türkiye

Received: 2024-01-10 **Accepted:** 2024-03-27 **Published Online:** 2024-03-29

Correspondence

Hüseyin Alper Kızıloğlu, MD, Assist. Prof.
Address: Tokat Gaziosmanpaşa University, Faculty of Medicine, Department of Radiology, Kaleardı Neighborhood, Muhittin Fisunoglu Street, 60250 Tokat, Türkiye
E-mail: alperkzloglu@hotmail.com

ABSTRACT

Objective: Breast cancer screening is a valuable field of health research conducted through mammography. However, mammography evaluation is the examination with the most frequent lack of to agrement among radiologists. In this study we aimed to show the compatibility of mammographic density classification with a new software, Bellus Breast Density Measurement Software (Option), with visual examination.

Methods: The mammographic density classification of 500 patients was retrospectively determined by five radiologists with varying levels of experience, according to the 5th version of the breast imaging reporting and data system (BIRADS). The mean age of 500 women included in the study was calculated as 53.8±10.08. The obtained data were compared with the Bellus software mammographic density classification of the same patients. Then, the visual evaluation and the compatibility of the Bellus software and the readers were compared.

Results: The agreement between the Bellus software and all five readers was poor (kappa value 0.07-0.12). The agreement of the readers with each other is moderate-good (kappa value 0.054-0.64). The Intraclass Correlation Coefficient (ICC) value for the five separate readers was calculated to be 0.80, indicating good compatibility, while the ICC value for the Bellus software with the five separate readers was calculated to be 0.74, indicating moderate compatibility. The Friedman test revealed that while the mammographic density classification of each reader remained consistent, the classification provided by the Bellus software differed.

Conclusion: Bellus Breast Density Measurement Software (Option) diagnostic accuracy is lower than visual examination. We recommend that the manufacturer develop the software.

Keyworlds Bellus software, mammographic density, automatic.

INTRODUCTION

Breast cancer is observed in high incidence and mortality for women in the world. According to the latest updated report of the American Cancer Society; breast cancer constitutes 30% of cancers in women and it is predicted that there will be 297,790 new cases in 2023. In addition, the incidence rate of breast cancer increases 0.5% each year [1]. For this reason, breast cancer screening programs play a vital role. Significant advances have recently occurred in the diagnosis and treatment of breast cancer, which has reduced mortality rates. Breast cancer screening is performed by mammography according to the breast imaging reporting and data system (BIRADS) all over the world. In our country, breast cancer screening is performed once a year for women over the age of 40. For patients in the high-risk group, screening is performed at an earlier age. These; family history, BRCA gene mutations, patients receiving radiotherapy to the chest area, and patients with high-risk pathology results such as atypia in previous breast biopsies [2]. In the presence of dense breasts in mammography, additional imaging is needed, and in this case, it is often evaluated with supplemental scanning ultrasonography [3,4].

Volumetric mammographic density (MD) defines the ratio of the percentage of dense tissue to the whole breast. On the image, it is related to the attenuation value of X-rays in the breast tissue [5]. Fat tissue appears radiolucent on mammography, which is dark on mammography. Fibroglandular tissue consists of fibroblasts, epithelial cells, and connective tissue cells, radiologically dense and brightly visible on mammography. According to BIRADS, breast density is divided into 4 major categories. These categories are defined as follows: Category A represents completely fatty tissue (5-24% fibroglandular tissue), B represents diffusely located fibroglandular tissue (25-49% fibroglandular tissue), C represents heterogeneous density (50-75% fibroglandular tissue), and D corresponds to dense breast density (75% and more fibroglandular tissue). In BIRADS version 5, categorization was made only by visual evaluation without including percentages in density categorization [6]. This classification is made visually by radiologists all over the world. Several software programs have been developed to automatically perform this subjective classification and quantify breast density. [7,8]. Dense breast definition defines categories C and D. Dense breast has two major clinical conditions. The first is that it reduces the sensitivity of mammographic screening and the second is that it is an independent risk factor for breast cancer [9,10]. As a result, dense breast tissue and cancerous tissue share similar attenuation characteristics, causing both to appear bright on mammography. Dense tissue and cancer tissue have similar attenuation and both appear bright on mammography.

Main Points:

- There is interobserver variability in determining mammographic density rates.
- Bellus software, which is automatic breast density calculation software, makes more errors than expected.
- In the Middle Black Sea Region, there are more cases in the nondense (A and B) mammographic density group.

Tomosynthesis or 3D mammography was approved by the US Food and Drug Administration (FDA) in 2011 for use in all clinical indications used for mammography. In this technique, the X-ray tube moves at a certain angle arc in the compressed breast tissue to obtain an image. In this way, images with smaller doses are obtained at multiple angles [11]. It has been shown that this method, especially in dense breasts, is more sensitive in detecting cancer compared to mammography [12,13].

Mammographic density classification is visually subjective, and differences in intraobserver and interobserver classification have been shown in the literature [14]. In addition, although it is not the standard method recommended for MD until now, visual assessment is performed with semi-automatic method and automatic methods. Studies in the literature compare these methods both with visual evaluations by radiologists and among themselves. [15,16]. These programs are; Cumulus (Sunnybrook Health Sciences Centre, Toronto, ON, Canada), Quantra version 2.0 (Hologic Inc, Bedford, MA, USA), Volpara Density Algorithm 1.5.0 (Volpara Health Technologies, Wellington, New Zealand), Densitas version 2.0.0 (Densitas Inc, Halifax, NS, Canada). However, there is currently no literature on the use of the Bellus automatic breast density measurement system specifically designed for tomosynthesis mammography devices developed by Fujifilm (Tokyo, Japan).

The aim of our study is to assess the agreement among observers in breast density classification using tomosynthesis, and to investigate the agreement among observers using the Bellus automatic breast density measurement system with the participation of different observers.

MATERIALS AND METHODS

The study was conducted retrospectively following approval from the local ethics committee (approval number: 22-KAEK-286, dated 22.12.2022). Our study adhered to the principles outlined in the Helsinki Declaration.

Case Selection

The mammography images of 650 randomly selected cases, who came to breast screening in accordance with national standards, were analyzed retrospectively. Patient selection was made randomly and is thought to objectively reflect the society. Cases who received hormone therapy before, cases with a history of breast cancer, cases who underwent breast surgery, cases with suspected malignancy in the breast during the examination, and cases with a difference in density between the two breasts were not included in the study. In addition, patients who could not perform adequate breast compression, had intense artifacts in the images, and patients in whom most of the breast was not included in the image area were not included in the study (Table 1). Ultimately, 500 patients were included in the study. The mean age of 500 women included in the study was calculated as 53.8±10.08.

Table 1. Mammographic Density classification rates of each reader and Bellus software.

Readers	A	B	C	D	Total
	n (%)	n (%)	n (%)	n (%)	n (%)
First	60 (12)	269 (53.8)	152 (30.4)	19 (3.8)	500 (100)
Second	68 (13.6)	196 (39.2)	182 (36.4)	54 (10.8)	500 (100)
Third	60 (12)	244 (48.8)	175 (35)	21 (4.2)	500 (100)
Fourth	64 (12.8)	224 (44.8)	175 (35)	37 (7.4)	500 (100)
Fifth	61 (12.2)	251 (50.2)	136 (27.2)	52 (10.4)	500 (100)
Bellus	210 (42)	245 (49)	42 (8.4)	3 (0.6)	500 (100)

n: number

Tomosynthesis Technique and Protocol

Images were obtained with the Fujifilm AMULET Innovality Digital Tomosynthesis device. The patients were given the maximum possible compression, mediolateral (MLO) and craniocaudal (CC). Due to automatic exposure control (AEC), milliampere-seconds (mAs) and kilovoltage peak (kVp) values differ from patient to patient. These values are in the range of 25-35 kVp and 25-100 mAs. Exposure time is less than 2 seconds and fully automatic exposure technique is used. Sections for tomosynthesis were created on the MLO. MLO images were obtained in such a way that the lower end of the pectoral muscle ends at the inferior level of the imaginary line passing through the areola and the inframammer fold is visible. CC images were obtained with the pectoral muscle visible in the posterior. No contrast agent was used.

Image Evaluation and Bellus Breast Density Measurement

The images were evaluated independently by 5 radiologists with different breast radiology experiences (Respectively, their experience was 2 years, 4 years, 9 years, 11 years, 16 years). One of the observers is a junior assistant and the other is a senior assistant. Three observers are radiologists. One of them is a breast radiologist and the other two are nonspecific radiologists. Observers have undergone international breast radiology training. Images were evaluated from a 21.3 inch 3MP IPS Screen medical monitor via Fujifilm Mammography Workstation (3000AWS7.0 Option) and Sectra IDS 7 PACS (Picture Archiving and Communication Systems). Readers performed the assessment according to BIRADS version 5. Breast densities of the patients were recorded as A, B, C, D (Figure 1). Observers made the BIRADS category in a non-quantitative way based on their own experience.

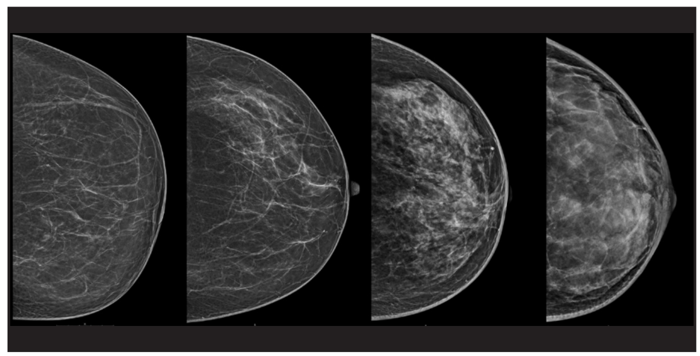

Figure 1. A-D mammographic density on CC radiographs of the left breast (left to right respectively).

Bellus Breast Density Measurement Software (Option) automatically categorized breast density into 4 groups using mammography images and exposure data. The calculation was categorized into 4 groups based on mammarian gland ratio (percentage), fat volume (cm^3) and total volume of the breast (cm^3) data, and threshold values. Bellus software calculates a percentage based on the ratio of fibroglandular tissue to fat tissue present in the image. This calculation is performed according to the density difference between fibroglandular tissue and fat tissue. The software performed these calculations on tomosynthesis data. These measurements were made by the software on 3D breast maps (Figure 2). In this categorization, the threshold values over the percentage of mammarian glands are; determined as 0-15% for group A, 15-35% for group B, 35-60% for group C and 60-100% for group D. These ranges are determined by the Bellus software.

Bellus Breast Density Measurement Software (Option) is a software belonging to Fujifilm, a Japanese company.

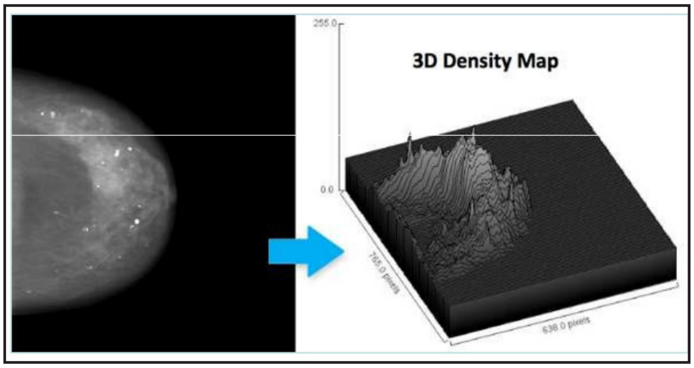

Figure 2. Schematic representation of the volumetric calculation of breast gland ratio (as a percentage), fat volume (cm^3) and breast volume (cm^3) by the Bellus Breast Density Measurement Software (Option) software developed by Fujifilm (Courtesy of Volpora Health).

Statistical Analysis

The SPSS 24 statistical software package (IBM Corp., Armonk, NY, USA) was used for all data analysis. Categorical measurements were summarized as numbers and percentages, and continuous measurements as mean, deviation, and minimum-maximum. It was checked with Friedman test whether 5 different readers and Bellus Breast Density Measurement Software (Option) median, Percentile 25 and 75 ratios were the same. Correlation between Cohen Kappa test and readers and Bellus Breast Density Measurement Software (Option) was examined. Kappa coefficient; if it is less than 0, weak agreement, 0-0.20 as insignificant agreement, 0.21-0.4 as low agreement, 0.41-0.6 as medium agreement, 0.61-0.8 as high agreement, and 0.81-1 as excellent agreement. The Intraclass Correlation Coefficient (ICC) value of 5 different readers and the Intraclass Correlation Coefficient value of the software with 5 different readers were calculated. Having this value below 0.5 is poor; moderate between 0.5-0.75; 0.75-0.9 good and above 0.9 were accepted as excellent agreement [17]. One-way analysis of variance (ANOVA) was used to compare the mean ages of each mammographic density group for each reader and software. Tukey HSD test was used for multiple comparison. P value <0.05 was considered statistically significant.

In a 6x4 design with 6 readers evaluating in 4 groups, approximately 500 samples will be studied with 80% power, 5% type I error and an effect size of 0.10. Sample size was calculated with the G*Power program (version 3.1.9.4).

RESULTS

All five readers included patients in the B, C, A and D categories, respectively, in mammographic density (From the highest number of patients in the groups to the lowest number). However, Bellus software included patients in the B, A, C and D categories, respectively, in mammographic density. (Table 2).

The dense breast category (groups C and D) was detected by the readers and the software at the following rates: first reader 171 (34.2%) patients, second reader 236 (47.2%) patients, third reader 196 (39.2%) patients, the fourth reader identified 212 (42.4%) patients, the fifth reader identified 188 (37.6%) patients, and the Bellus software identified 45 (9%) patients. The rate of nondense breasts was higher for all five readers and Bellus software in the patients included in the study. Dense breast percentage was determined most in the second reader and least in Bellus software.

In the Friedman test, it was determined that the median, Percentile 25 and 75 ratios were the same for five different readers. For Bellus software, the median value is the same, but the Percentile 25 and 75 values are different (Table 3).

In the visual evaluation, the agreement of the readers with each other was calculated as medium and high. In the agreement of the readers with each other, the highest agreement kappa value of 0.66 is between the second and fourth readers, and the lowest agreement kappa value of 0.54 is between the first and second readers. The agreement of each reader with the Bellus software was insignificant. With Bellus software, the highest agreement kappa coefficient was found between 0.12 and the first reader, and the lowest agreement was between 0.04 and the second reader (Tables 4 and 5).

Visually, the Intraclass Correlation Coefficient (ICC) value was calculated as 0.8 (good correlation) for 5 readers and 0.74 (moderate correlation) when Bellus software was included with 5 readers (Tables 4 and 5).

The mean age of 500 women included in the study was calculated as 53.8±10.08. Except for the second reader, no statistically significant difference was found for the C and D groups in terms of mean age in the other four readers (p<0.001). However, no statistical significance was found for any group in the mean age of the second reader. In Bellus software, no statistically significant difference was found for both groups C and D, and groups B and D (p<0.001). In each reader and Bellus software, the highest mean age was determined in group A, and the mean age was found in groups B, C, and D with decreasing mean ages, respectively (Table 6).

The results of the power analysis were determined as 80% power, 5% type I error and an effect size of 0.10.

Table 2. Median, percentile 25 and 75 values of each reader and Bellus software in Friedman test.

Readers	Median	Percentile 25	Percentile 75
First	2	2	3
Second	2	2	3
Third	2	2	3
Fourth	2	2	3
Fifth	2	2	3
Bellus	2	1	2

Table 3. Inter-reader agreement

Inter-reader agreement	Kappa value	Agreement Level	ICC value(95%)
First and Second	0.54	medium agreement	0.80 (good)
First and Third	0.63	high agreement	
First and Fourth	0.59	medium agreement	
First and Fifth	0.59	medium agreement	
Second and Third	0.64	high agreement	
Second and Fourth	0.66	high agreement	
Second and Fifth	0.63	high agreement	
Third and Fourth	0.62	high agreement	
Third and Fifth	0.60	medium agreement	
Fourth and Fifth	0.64	high agreement	

Abbreviations: ICC: Intraclass Correlation Coefficient

Table 4. Agreement between readers and Bellus software

Bellus Software	Kappa value	Agreement Level	ICC value (95%)
First	0.12	insignificant agreement	0.74 (moderate)
Second	0.04	insignificant agreement	
Third	0.07	insignificant agreement	
Fourth	0.07	insignificant agreement	
Fifth	0.09	insignificant agreement	

Abbreviations: ICC: Intraclass Correlation Coefficient

Table 5. Average age in Mammographic Density groups of each reader and Bellus software

Readers	Total	Age			
		A	B	C	D
	Average ±SD	Average ±SD	Average ±SD	Average ±SD	Average ±SD
First	53.8±10.08	61.42±9.01	56.21±9.83	47.68±7	44.68±4.28
Second	53.8±10.08	61.46±8.78	56.88±9.64	49.91±8.75	46.13±5.78
Third	53.8±10.08	62.57±8.62	55.96±9.88	48.95±7.83	44.1±2.53
Fourth	53.8±10.08	61.83±8.75	56.57±9.85	48.96±7.85	46.08±6.19
Fifth	53.8±10.08	61.57±9.81	56.15±9.99	48.94±7.09	46.1±5.98
Bellus	53.8±10.08	57.2±9.18	52.51±10.3	45.21±4.67	42.67±4.73

Abbreviations: SD: Standard Deviation

DISCUSSION

We have demonstrated in our study that the new software, Bellus automatic breast density measurement, which assesses through tomosynthesis, exhibits lower accuracy in evaluating breast density compared to visual examination, as determined by multiple readers with varying levels of experience. Additionally, we concluded that there is a good level of agreement among radiologists with different levels of experience who classify breast density through visual examination. Furthermore, we found that the youngest group of breast density patterns is group D, while the oldest group is group A, as identified by both visual radiologists and the Bellus software program.

Our study aims to evaluate a novel software for classifying mammographic density through tomosynthesis, a methodology not previously explored. While similar studies have been conducted on single-section mammography images, ours marks the first attempt using tomosynthesis. Furthermore, the involvement of five radiologists with varying levels of experience and the examination of a large number of samples enhance the significance of our study. By employing these methods, we aimed to objectively evaluate mammographic density, a concept previously challenging to classify accurately.

There are numerous studies in the literature that incorporate volumetric estimates for BIRADS. Brandt et al. conducted a study comparing automatic breast density measurement with clinical measurement, revealing medium agreement between Volpara and Quanta software and clinical breast density classification (kappa values: 0.57 and 0.46, respectively) [18]. Another study comparing visual evaluation with Volpara and Quanta software in the 5th version of BIRADS found low to medium agreement (kappa value: 0.32-0.43) and medium to high agreement (kappa value: 0.54-0.61) [19]. When comparing breast density studies with BIRADS version 4, Volpara software showed low to high agreement (kappa value: 0.4-0.8) and Quanta software showed high agreement (kappa value: 0.63-0.73) [20-23]. In our study, we compared Bellus, a newer software, and found insignificant agreement for each reader (kappa value: 0.04-0.12). Additionally, we observed differences in Percentile 25 and 75 values, although the median value was the same among five observers in the Friedman test. We attribute the lower level of agreement in our study to the novelty of the software we evaluated.

In the literature, mammographic density percentages have been extensively studied for BIRADS version 5. Previous studies have reported varying percentages across different density categories. For instance, one study found that 10% of patients were classified as fatty (group A), 40% as scattered (group B), 40% as heterogeneously dense (group C), and 10% as extremely dense (group D) [24,25]. Another study reported percentages of 3.5%, 22.1%, 54.9%, and 19.5% for groups A, B, C, and D, respectively [19]. Similarly, another study found rates of 1.6%, 14.3%, 69.1%, and 15% in groups A, B, C, and D, respectively [26].

In our study, we observed differences in the rates reported by five observers, excluding the Bellus software, with percentages ranging from 12% to 13.6% for group A, 39.2% to 53.8% for group B, 27.2% to 36.4% for group C, and 3.8% to 10.8% for group D. Initially, we attributed these discrepancies to variations in the studied population. Contrary to previous studies, our study, conducted in the Central Black Sea Region, found a higher rate of non-dense breasts in our study population. This suggests that mammographic screening may be more sensitive in our society, potentially reducing the need for supplemental screening.

In the literature, the agreement between observers in mammographic density classification has been extensively evaluated, revealing wide variability ranging from poor to perfect agreement (kappa value: 0.02-0.87) [21,26,27]. Some studies reported better interobserver agreement, reaching perfect agreement with kappa values ranging from 0.81 to 0.84.

In our study, we found moderate to high inter-reader agreement (kappa value: 0.54-0.66), with an Intraclass Correlation Coefficient (ICC) value of 0.8 indicating good correlation among readers. We attributed the differences observed in the literature to the varying levels of experience among radiologists. Previous studies have linked inter- and intra-reader variability in mammographic density determination to the differing experiences of radiologists [28].

More objective presentation of mammographic density classification, which is a subjective evaluation, with software will determine the needs of patients for Supplemental Screening more accurately. It will also assist the radiologists in their process of gaining experience.
We believe that the reason the Bellus software provides lower estimates of breast density rates compared to human radiologists is due to its novelty. As the software is still in the testing phase, our initial study highlights the need for improvements. While radiologists typically rely on qualitative assessments in determining breast density rates through their daily practice, software applications like Bellus aim to incorporate more quantitative data into their findings.

Limitations

The limitations of our study include the following: it was a single-center analysis, cases with different mammographic densities of both breasts were not included, and the behavior of the Bellus software in such cases was not tested. Additionally, the rate of dense breasts in our study was low, and we did not evaluate intraobserver agreement. Furthermore, we did not compare the Bellus software program with more common software programs. In the tomosynthesis technique, variations in kVp and mAs values from patient to patient may affect automatic density calculation. Although we do not believe this situation significantly impacts our results, it should be noted as a limitation.

CONCLUSION

A new software program, Bellus Breast Density Measurement Software [Optional function of AMULET Innovality (3000AWS7.0 Option)], has lower diagnostic accuracy than visual examination. We recommend that the manufacturer develop the software. In addition, the relatively higher rate of detection of patients in the nondense mammographic density group in the region we live in (Central Black Sea Region) suggests that there will be a lesser need for Supplemental Screening in this region.

Acknowledgement
We would like to thank Volpora Health for permission to use the schematic image in Figure 2.

Conflict of Interest: The authors declare that there is no conflict of interest.

Funding: No source of funding.

Author Contributions: Conception: HAK; Design: MB; Supervision: EG, Materials: MFB, MEÇ, YB; Data Collection and/or Processing: HAK, MB, EG, YB, MFB, MEÇ; Analysis and/or Interpretation: OD; Literature: HAK, Review: HAK, MB, EG, YB, MFB, MEÇ; Writing: HAK, MB, EG, YB, MFB, MEÇ; Critical; Review: HAK, MB, EG, YB, MFB, MEÇ.

Ethical Approval: This study was approved by the local ethical committee of the city hospital (202/22-KAEK-286) and the procedures were according to the ethical standards of the responsible committee on human experimentation.

Author Contributions: All the authors undertook the literature review, planned the study, wrote, and reviewed the manuscript.

REFERENCES

[1] American Cancer Society. Cancer Facts and Figures 2023. Atlanta, Ga: American Cancer Society. 2023.

[2] Smith RA, Saslow D, Sawyer KA, Burke W, Costanza ME et al. American Cancer Society High-Risk Work Group; American Cancer Society Screening Older Women Work

Group; American Cancer Society Mammography Work Group; American Cancer Society Physical Examination Work Group; American Cancer Society New Technologies Work Group; American Cancer Society Breast Cancer Advisory Group. American Cancer Society guidelines for breast cancer screening: update 2003. CA Cancer J Clin. 2003;53(3):141-169. https://doi.org/10.3322/canjclin.53.3.141

[3] Scheel JR, Lee JM, Sprague BL, Lee CI, Lehman CD. Screening ultrasound as an adjunct to mammography in women with mammographically dense breasts. Am J Obstet Gynecol. 2015; 212: 9-17. https://doi.org/10.1016/j.ajog.2014.06.048

[4] Sprague BL, Stout NK, Schechter C, van Ravesteyn NT, Cevik M et al. Benefits, harms, and cost-effectiveness of supplemental ultrasonography screening for women with dense breasts. Ann Intern Med. 2015;62:157-166. https://doi.org/10.7326/M14-0692

[5] Colin C, Schott AM. Re: Breast tissue composition and susceptibility to breast cancer. J Natl Cancer Inst. 2011;103(1):77. https://doi.org/10.1093/jnci/djq464

[6] Spak DA, Plaxco JS, Santiago L, Dryden MJ, Dogan BE. BI-RADS® fifth edition: A summary of changes. Diagn Interv Imaging. 2017;98(3):179-190. https://doi.org/10.1016/j.diii.2017.01.001.

[7] Jeffreys M, Harvey J, Highnam R. Comparing a new volumetric breast density method (VolparaTM) to cumulus. In: Digital mammography: 2010/2010. Berlin: Springer; 2010.pp408-413. https://doi.org/10.1007/978-3-642-13666-5_55

[8] Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. Automated analysis of mammographic densities. Phys Med Biol 1996;41:909-923. https://doi.org/10.1088/0031-9155/41/5/007

[9] Kerlikowske K, Scott CG, Mahmoudzadeh AP, Ma L, Winham S et al. Automated and Clinical Breast Imaging Reporting and Data System Density Measures Predict Risk for Screen-Detected and Interval Cancers: A Case-Control Study. Ann Intern Med. 2018;168(11):757-765. https://doi.org/10.7326/M17-3008.

[10] Puliti D, Zappa M, Giorgi Rossi P, Pierpaoli E, Manneschi G et al. Volumetric breast density and risk of advanced cancers after a negative screening episode: a cohort study. Breast Cancer Res. 2018;20(1):95. https://doi.org/10.1186/s13058-018-1025-8

[11] Baker JA, Lo JY. Breast tomosynthesis: state-of-the-art and review of the literature. Acad Radiol. 2011;18(10):1298-1310. https://doi.org/10.1016/j.acra.2011.06.011

[12] Skaane P, Bandos AI, Niklason LT, Sebuødegård S, Østerås BH et al. Digital Mammography versus Digital Mammography Plus Tomosynthesis in Breast Cancer Screening: The Oslo Tomosynthesis Screening Trial. Radiology. 2019:182394. https://doi.org/10.1148/radiol.2019182394

[13] Pattacini P, Nitrosi A, Giorgi Rossi P, Iotti V, Ginocchi V et al. Digital Mammography versus Digital Mammography Plus Tomosynthesis for Breast Cancer Screening: The Reggio Emilia Tomosynthesis Randomized Trial. Radiology. 2018;288(2):375-385. https://doi.org/10.1148/radiol.2018172119

[14] Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B et al. Categorizing breast mammographic density: intra- and interobserver reproducibility of BI-RADS density categories. Breast. 2005;14:269-275. https://doi.org/10.1016/j.breast.2004.12.004

[15] Jeffreys M, Harvey J, Highnam R. Comparing a New Volumetric Breast Density Method (Volpara) to Cumulus. In: Martí J, Oliver A, Freixenet J, Martí R, editors. Lecture Notes in Computer Science: 10th International Workshop on Digital Mammography; 2010 Jun 16–18; Girona, Spain: Springer-Verlag; 2010.p.408-413. https://doi.org/10.1007/978-3-642

[16] Regini E, Mariscotti G, Durando M, Ghione G, Luparia A et al. Radiological assessment of breast density by visual classification (BI-RADS) compared to automated volumetric digital software (Quantra): implications for clinical practice. Radiol Med. 2014;119:741-749. https://doi.org/10.1007/s11547-014-0390-3

[17] Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. Journal of Chiropractic Medicine. 2016;15(2):155-163. https://doi.org/10.1016/j.jcm.2016.02.012

[18] Brandt KR, Scott CG, Ma L, Mahmoudzadeh AP, Jensen

MR et al. Comparison of Clinical and Automated Breast Density Measurements: Implications for Risk Prediction and Supplemental Screening. Radiology. 2016;279(3):710-719. https://doi.org/10.1148/radiol.2015151261

[19] Youk JH, Gweon HM, Son EJ, Kim JA. Automated Volumetric Breast Density Measurements in the Era of the BI-RADS Fifth Edition: A Comparison With Visual Assessment. AJR Am J Roentgenol. 2016;206(5):1056-1062. https://doi.org/10.2214/AJR.15.15472

[20] Magni V, Interlenghi M, Cozzi A, Alì M, Salvatore C et al. Development and Validation of an AI-driven Mammographic Breast Density Classification Tool Based on Radiologist Consensus. Radiology: Artificial Intelligence. 4(2):e210199. https://doi.org/10.1148/ryai.210199

[21] Yoshida R, Yamauchi T, Akashi-Tanaka S, Matsuyanagi M, Taruno K et al. Optimal Breast Density Characterization Using a Three-Dimensional Automated Breast Densitometry System. Current Oncology. 2021;28(6):5384-5394. https://doi.org/10.3390/curroncol28060448

[22] Dontchos BN, Yala A, Barzilay R, Xiang J, Lehman CD. External Validation of a Deep Learning Model for Predicting Mammographic Breast Density in Routine Clinical Practice. Academic Radiology. 2021;28(4):475-480. https://doi.org/10.1016/j.acra.2019.12.012

[23] Alomaim W, O'Leary D, Ryan J, Rainford L, Evanoff M, et al. Variability of Breast Density Classification Between US and UK Radiologists. Journal of Medical Imaging and Radiation Sciences. 2019;50(1):53-61. https://doi.org/10.1016/j.jmir.2018.11.002

[24] Alomaim W, O'Leary D, Ryan J, Rainford L, Evanoff M, et al. Subjective Versus Quantitative Methods of Assessing Breast Density. Diagnostics. 2020;10(5):331. https://doi.org/10.3390/diagnostics10050331

[25] Destounis SV, Santacroce A, Arieno A. Update on Breast Density, Risk Estimation, and Supplemental Screening. American Journal of Roentgenology. 2020;214(2):296-305. https://doi.org/10.2214/AJR.19.21994

[26] Li H, Mukundan R, Boyd S. Breast Density Classification Using Multifractal Spectrum with Histogram Analysis. 2019 International Conference on Image and Vision Computing New Zealand (IVCNZ), Dunedin, New Zealand. 2019;1-6. http://doi.org/10.1109/IVCNZ48456.2019.8961037

[27] Balleyguier C, Arfi-Rouche J, Boyer B, Gauthier E, Helin V et al. A new automated method to evaluate 2D mammographic breast density according to BI-RADS® Atlas Fifth Edition recommendations. European Radiology. 2019;29(7):3830-3838. https://doi.org/10.1007/s00330-019-06016-y

[28] Ciatto S, Houssami N, Apruzzese A, Bassetti E, Brancato B et al. Reader variability in reporting breast imaging according to BI-RADS assessment categories (the Florence experience). Breast. 2006;15:44-51. https://doi.org/10.1016/j.breast.2005.04.019

How to Cite;

Kızıloğlu HA Beyhan M, Gökçe E, Birişik Y, Battal MF, Çeker, ME Demir O (2024) Evaluation of Automated Mammographic Density Classification in Tomosynthesis: Comparison with Radiologists. Eur J Ther. https://doi.org/10.58600/eurjther2002