Radiologist Assessment of Breast Density by BI-RADS Categories Versus Fully Automated Volumetric Assessment

Hye Mi Gweon[1]
Ji Hyun Youk
Jeong-Ah Kim
Eun Ju Son

OBJECTIVE. The objective of our study was to estimate mammographic breast density using a fully automated volumetric breast density measurement method in comparison with BI-RADS breast density categories determined by radiologists.

MATERIALS AND METHODS. A total of 791 full-field digital mammography examinations with standard views were evaluated by three blinded radiologists as BI-RADS density categories 1–4. For fully automated volumetric analysis, volumetric breast density was calculated with fully automated software. The volume of fibroglandular tissue, the volume of the breast, and the volumetric percentage density were provided.

RESULTS. The weighted overall kappa was 0.48 (moderate agreement) for the three radiologists' estimates of BI-RADS density. Pairwise comparisons of the radiologists' measurements of BI-RADS density revealed moderate to substantial agreement, with kappa values ranging from 0.51 to 0.64. There was a significant difference in mean volumetric breast density among the BI-RADS density categories, and the mean volumetric breast density increased as the BI-RADS density category increased ($p < 0.001$). A significant positive correlation was found between BI-RADS categories and fully automated volumetric breast density ($\rho = 0.765$, $p < 0.001$).

CONCLUSION. Our study showed good correlation of the fully automated volumetric method with radiologist-assigned BI-RADS density categories. Mammographic density assessment with the fully automated volumetric method may be used to assign BI-RADS density categories.

BI-RADS is routinely used in clinical practice for reporting mammographic breast density [1]. In previous studies, however, researchers found that the assignment of density categories showed varying degrees of agreement among radiologists (κ = 0.43–0.76) [2–4]. Thus, there is increasing interest in the use of quantitative measures derived from mammographic images. Quantitative measures of breast composition can be derived from a variety of visual computer-assisted and fully automated methods [4–6]. There are two methods of breast composition measures: area and volumetric. The area-based method works directly from the image of the compressed and projected breast and attempts to quantify breast density by segmenting areas of the mammographic images [7, 8]. However, this method is intrinsically subjective and can be limited by substantial inter- and intraobserver variability. Also, it requires additional decision time even by skilled users [4, 9].

The limitations of the area-based method have led several research groups to develop volumetric measures of breast density. The volumetric method quantifies breast density by using the pixel value to derive information about the x-ray attenuation properties of the column of breast tissue above that pixel [10, 11]. In recent years, a fully automated method of measuring breast volume to estimate breast density has been developed [11, 12]. Its advantages include the elimination of user variability, elimination of time-consuming density estimation, and consideration of the breast as a 3D organ. However, there are few studies of fully automated volumetric measurements of breast density in clinical practice [12, 13] and to date there is no study that compares the measurements of qualitative and quantitative methods—that is, radiologist-assigned BI-RADS category and a fully automated volumetric method. Therefore, the purpose of our study was to estimate mammographic breast density using a fully automated volumetric breast density measurement

Keywords: automated evaluation, breast density, breast imaging, mammography

DOI:10.2214/AJR.12.10197

Received October 22, 2012; accepted after revision December 20, 2012.

[1]All authors: Department of Radiology, Gangnam Severance Hospital, Yonsei University College of Medicine, 211 Eonju-ro, Dogok-Dong, Gangnam-Gu, Seoul 135-720, South Korea. Address correspondence to J. H. Youk (jhyouk@yuhs.ac).

AJR 2013; 201:692–697

0361–803X/13/2013–692

method in comparison with BI-RADS breast density categories determined by radiologists.

Materials and Methods

Patients and Mammography

From February 2012 to March 2012, 791 full-field digital mammography (FFDM) examinations with standard views were performed at our institution. Standard craniocaudal and mediolateral oblique views made up the dataset in this study.

Digital mammographic examinations were performed on an FFDM unit (Senographe 2000D, GE Healthcare; or Lorad Selenia, Hologic). The Senographe 2000D has a 100-μm digital spatial resolution and produces an image data matrix size of 1914 × 2294 pixels. The Lorad Selenia unit is equipped with 24 × 29 cm amorphous selenium detectors with a pixel size of 70 μm. The two different digital mammography systems were randomly assigned to each patient according to the mammography suite schedule.

BI-RADS Density Classification

All mammographic images were downloaded to a soft-copy review workstation (Selenia Softcopy Workstation, Hologic) with soft-copy reading software (MeVis BreastCare version 6.0.5, MeVis Medical Solutions). Three blinded radiologists who specialize in breast imaging and at the time of the study had 5–10 years of experience in interpreting mammography and 5–8 years of experience in soft-copy review of digital mammography independently reviewed the images at the review workstation. Each mammogram was assessed for breast density according to the BI-RADS breast density categories. The following BI-RADS categories for breast density [1] were used for mammographic interpretations: category 1, almost fatty (< 25% glandular); category 2, scattered fibroglandular densities (25–50% glandular); category 3, heterogeneously dense (51–75% glandular); and category 4, extremely dense (> 75% glandular).

Mammographic Density Analysis by Fully Automated Volumetric Software

For fully automated volumetric analysis, Volpara software (version 1.5.1, Mātakina Technology) was used. By using the DICOM for processing image data generated by the digital mammography system, the Volpara algorithm calculates an objective measurement of breast density using volumetric parameters. The first step is to find a reference point in the breast that is of known breast composition. Typically this point would be an area near the chest wall of all fatty tissue. The second step is to compute the x-ray attenuation at each pixel and then to convert that attenuation to an estimate of the tissue composition between that pixel and the x-ray source. That step creates a density map. By adding up all the values in the density map, the software can compute the volume of fibroglandular tissue in cubic centimeters, the volume of breast in cubic centimeters, and the percentage of the breast that consists of fibroglandular tissue. The software uses these data to determine the volumetric density. The volumetric density ranges from 0% to 40%. The breast density information is provided per breast (by averaging the craniocaudal and mediolateral oblique values). For each patient, a Volpara density grade (VDG) is also provided. The VDG is the result of mapping the average volumetric breast density for the patient corresponding to BI-RADS breast density category determined by using an expert opinion from the University of Virginia. VDG was graded as follows: 0–4.7% volumetric density, VDG 1; 4.8–7.9%, VDG 2; 8.0–15.0%, VDG 3; and 15.1% and up, VDG 4. Once Volpara recognizes that it has a complete study, it automatically begins processing data and is able to send out DICOM secondary capture images with the breast density information (Fig. 1).

Data and Statistical Analysis

Each patient's medical record was reviewed and data about age and personal history of breast augmentation, breast-conserving surgery, or mastectomy were compiled. Also, radiologic reports and mammograms were reviewed for any mass or calcification larger than 1 mm.

Weighted kappa values were calculated to assess the proportion of interobserver agreement about breast density according to BI-RADS category. The kappa values were interpreted as suggested by Landis and Koch [14] as follows: A kappa value equal to or less than 0.20 indicates slight agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, substantial agreement; and 0.81–1.00, almost perfect agreement. Statistical comparisons were performed with independent and paired Student *t* tests or with the analysis of variance with Bonferroni adjustment for multiple comparisons for continuous variables and the chi-square or Fisher exact test for categoric variables. The correlation between the BI-RADS density category and volumetric breast density provided by the fully automated software was estimated using the Spearman rank correlation coefficient (ρ). Statistical analysis was performed using statistics software (SAS, version 9.1.3, SAS Institute). Differences were considered to be statistically significant at $p < 0.05$.

Results

Of the 791 FFDM examinations, 778 examinations of 778 patients 30–86 years old (mean age, 51.7 years) were enrolled in this study. The remaining 13 examinations (1.6%), which we excluded from the study, were examinations of breasts augmented with foreign material injection ($n = 2$) or saline bag insertion ($n = 10$) and an examination in which the software failed to obtain the data (i.e., volume of breast, volume of fibroglandular tissue, and breast density in DICOM capture image) because of technical error ($n = 1$).

Among the 778 patients, 178 (22.9%) had a history of breast cancer surgery (unilateral mastectomy in 115 and breast-conserving surgery in 63). A review of the patients' mammograms showed that 44 patients (5.6%) had abnormal mammographic findings: masses in 27 and calcifications in 17. In cases in which the mass or calcification was visible on the mammograms, no significant difference was found in the mean volumetric breast density between the abnormal breast and contralateral normal breast (mean volumetric density of normal breast vs breast with a mass: 16.4% ± 7.1% [SD] vs 16.1% ± 7.8%, $p = 0.74$; mean volumetric density of normal breast vs breast with a calcifica-

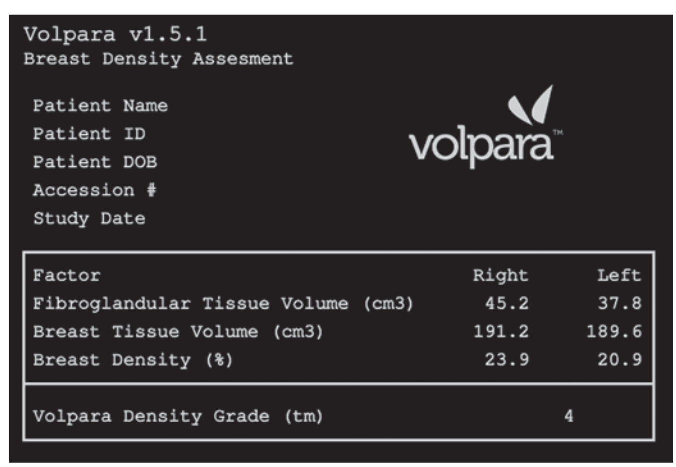

Fig. 1—Example of breast density assessment made using fully automated volumetric method (Volpara software, version 1.5.1, Mātakina Technology). ID = identification, DOB = date of birth.

tion: 21.0% ± 10.2% vs 20.9% ± 10.1%, p = 0.84). For the breasts iatrogenically altered by cancer surgery, however, the mean volumetric breast density of the altered breast (16.7% ± 7.3%) was significantly higher than that of the contralateral normal breast (14.0% ± 7.8%) ($p < 0.001$). Because of these results, mammograms of only the normal breasts of the 178 patients who underwent breast cancer surgery were included in this study and the VDG was reassessed according to the volumetric breast density of the normal breast only.

Regarding the two different mammography units, 346 mammographic examinations (44.5%) were performed using the Senographe 2000D and the remaining 432 (55.5%) were acquired using the Lorad Selenia. There were no significant differences in patient age, data from the automated software, and BI-RADS breast density category between the two different systems (Table 1).

The overall weighted kappa of the three radiologists' estimates of BI-RADS density categories showed moderate agreement (κ = 0.48). Pairwise estimates of the weighted kappa between two different observers showed moderate to substantial agreement (κ = 0.51–0.64). Table 2 summarizes the results of each observer's estimate of BI-RADS density category and VDG. There were no significant differences for BI-RADS density category among the three observers (p = 0.70). However, a significant difference was found among all observers' estimates and VDGs (p = 0.001). Pairwise estimates of the weighted kappa between BI-RADS density category by two radiologists' agreement and VDG showed moderate agreement (κ = 0.54) (Table 3). According to the results from overall observers' estimates of BI-RADS density category, BI-RADS category 3 was found in 61.9% of the cases and BI-RADS category 4, in 20.2%. Regarding the automated results from the Volpara software, VDG 3 was found in 44.7% and VDG 4 in 41.6%.

Of the 778 mammographic examinations, 497 examinations (63.9%) showed agreement on BI-RADS density category among the three observers and all examinations (100%) showed agreement on BI-RADS density category among at least two observers. Even in cases of disagreement about the BI-RADS density category, the differences were within one category. Table 4 shows breast density data yielded by the fully automated volumetric method for all cases according to BI-RADS density categories based on agreement of at least two radiologists. There was a significant difference in mean volumetric breast density among BI-RADS density categories and mean volumetric breast density increased as BI-RADS density category increased (p < 0.001; after Bonferroni adjustment for multiple comparisons, p < 0.001 except BI-RADS category 1 vs 2 [p > 0.99]) (Fig. 2A). A significant positive correlation was found between BI-RADS categories and fully automated volumetric breast density (ρ = 0.765 overall, ρ = 0.785 for Senographe 2000D, ρ = 0.745 for Lorad Selenia; p < 0.001). Regarding the 497 cases in which agreement of BI-RADS density categories was shown among all three radiologists, there was also a significant difference in mean volumetric breast density among BI-RADS density categories and mean volumetric breast density increased as BI-RADS density category increased (p < 0.001 overall; after Bonferroni adjustment for multiple comparisons, p < 0.001 except BI-RADS category 1 vs 2 [p = 0.14]). A significant positive correlation was also found between BI-RADS categories and fully automated volumetric breast density (ρ = 0.744 overall, ρ = 0.749 for Senographe 2000D, ρ = 0.757 for Lorad Selenia; p < 0.001) (Fig. 2B).

Discussion

Mammographic breast density is considered to be an important independent risk factor for breast cancer [15–19]. The relationship

TABLE 1: Patient Age and Breast Density Results Overall and by Digital Mammography System

Characteristic or Result	All Examinations (n = 778)	Senographe 2000D[a] (n = 346)	Lorad Selenia[b] (n = 432)	p[c]
Patient age (y), mean ± SD	51.7 ± 9.2	51.9 ± 9.6	51.6 ± 8.9	0.75
Volpara[d] results, mean ± SD				
Fibroglandular tissue volume (cm^3)	51.0 ± 25.7	49.7 ± 39.6	52.0 ± 28.3	0.71
Breast volume (cm^3)	380.9 ± 204.9	379.2 ± 189.3	383.3 ± 225.5	0.20
Breast density (%)	15.4 ± 7.8	15.2 ± 8.4	15.8 ± 8.2	0.35
VDG, no. (%) of patients				0.54
1	4 (0.5)	1 (0.3)	3 (0.7)	
2	102 (13.1)	50 (14.5)	52 (12.0)	
3	348 (44.7)	157 (45.4)	191 (44.2)	
4	324 (41.6)	138 (39.9)	186 (43.1)	
Radiologists' results[e]				
BI-RADS density category, no. (%) of patients				0.47
1	19 (2.4)	8 (2.3)	11 (2.5)	
2	131 (16.8)	59 (17.1)	72 (16.7)	
3	482 (62.0)	206 (59.5)	276 (63.9)	
4	146 (18.8)	73 (21.1)	73 (16.9)	

Note—VDG = Volpara density grade.
[a]GE Healthcare.
[b]Hologic.
[c]The p values indicate comparison between Senographe 2000D and Lorad Selenia.
[d]Version 1.5.1, Mātakina Technology.
[e]At least two radiologists agreed about the BI-RADS category in all cases.

TABLE 2: Frequency of Breast Mammographic Density Grades Using Four-Grade Scale

Breast Density Grade[a]	No. (%) of Mammographic Examinations				
	Observers				Volpara Software[b]
	Observer A	Observer B	Observer C	Overall	
1	20 (2.6)	5 (0.6)	20 (2.6)	45 (1.9)	4 (0.5)
2	141 (18.1)	110 (14.1)	121 (15.6)	372 (15.9)	102 (13.1)
3	457 (58.7)	517 (66.5)	471 (60.5)	1445 (61.9)	348 (44.7)
4	160 (20.6)	146 (18.8)	166 (21.3)	472 (20.2)	324 (41.6)

[a]Observers assessed breast density according to BI-RADS density categories 1–4, whereas Volpara software assigned density grades, which it refers to as "Volpara density grade" or "VDG," 1–4.
[b]Version 1.5.1, Mātakina Technology.

TABLE 3: Frequency of Volpara[a] Density Grades (VDGs) for Each BI-RADS Category Assigned by Agreement of at Least Two Radiologists

BI-RADS Category	No. (%) of Mammographic Examinations ($n = 778$)				Total No. of Examinations
	VDG 1	VDG 2	VDG 3	VDG 4	
1	3 (15.8)	12 (63.2)	4 (21.1)	0 (0)	19
2	1 (0.8)	79 (60.3)	51 (38.9)	0 (0)	131
3	0 (0)	11 (2.3)	285 (59.1)	186 (38.6)	482
4	0 (0)	0 (0)	8 (5.5)	138 (94.5)	146

[a]Version 1.5.1, Mātakina Technology.

between increased breast density and cancer risk is important in clinical practice, especially for breast cancer risk prediction [20]. Without an accurate estimation, breast density is unlikely to become useful in risk assessment models and clinical trials [21]. Measurement reproducibility of breast density is extremely important for cancer risk predictive models and epidemiologic studies. The method most commonly used to classify mammography density is BI-RADS; however, classification is commonly determined on a visual basis, so it is subjective and has been associated with suboptimal reproducibility [2–4]. The suboptimal reproducibility of this method could affect the ability of the BI-RADS density categories, even when incorporated into a larger model, to predict breast cancer risk. In our study, interobserver agreement in reporting breast density according to BI-RADS categories, moderate to substantial, was not satisfactory, which is compatible with the findings of previous studies [2–4].

Breast density evaluation with continuous and more precise values than the BI-RADS categories is a clinical need that prompted the development of software for quantitative assessment of breast density on standard mammograms. One of the most frequently used techniques for the quantification of breast density on a mammogram is a computer-assisted threshold method [7]. A limitation of this method is that it does not take into account the thickness of the dense tissue. Moreover, technical characteristics such as radiation dose and projection angle are seldom registered but can influence the density measurement. Another disadvantage is the subjectivity introduced by the reader who sets the threshold manually, which could cause observer bias. The idea of automated volumetric assessment of breast density is definitely appealing because it is supposed to have absolute reproducibility. On the other hand, recently developed volumetric methods for FFDM can take into account different energy spectra for different anode target and filter materials, information that is automatically stored in the header of raw FFDM data. The calibration data, to be used for the volumetric estimations, from FFDM images are more reliable than those that can be obtained from film-screen mammograms. In addition, FFDM has the advantage that the pixel value is known to be linearly related to exposure. As a result, estimates of volumetric breast density are expected to be more accurate and reproducible for FFDM than for film-screen mammography. However, considering that all we know about breast density as a risk factor or as a determinant of mammography sensitivity is based on visual classification, breast density values obtained by automated software should be adapted to match commonly used BI-RADS categories. Our results showed that volumetric breast density evaluated by the fully automated volumetric method correlated well with BI-RADS categories. Tagliafico et al. [22] reported that there was good correlation between BI-RADS categories and breast density evaluated with fully automated software (area-based density estimates) ($r = 0.62$, $p < 0.01$). Also, we found a significant difference in mean volumetric breast density among the BI-RADS density categories (Table 4).

In detail, however, the difference in volumetric breast density between BI-RADS categories 1 and 2 was not significant after adjustment for multiple comparisons. Also, volumetric breast densities ranged widely for single BI-RADS categories and overlapped with different BI-RADS categories, especially BI-RADS categories 3 and 4. For example, BI-RADS category 3 included mammograms with volumetric breast densities that ranged from 5.7% to 32.5% (Fig. 2). In a study about a comparison of qualitative estimates based on BI-RADS categories and the quantitative determinations of density made using computer-aided mammographic density estimates [4], similar results were found that breast densities by human visual assessment could be classified into the upper or lower categories when the breast density difference is low. The use of the automated software could eliminate this problem because the density measurement gives a continuous scale.

When a four-grade scale of volumetric breast density was kept in fully automated volumetric method (VDGs 1–4) and compared with the BI-RADS (categories 1–4), a significant difference was found (Table 2). The fully automated volumetric software classified more mammographic examinations as density grade 4 (41.6%) than the three observers (20.2%) (Table 2). This overestimation can be reduced by means of a correction in the mapping between average volumetric breast density and VDGs 1–4.

In this study, two different systems were used, Senographe 2000D and Lorad Selenia. In previous studies [23–25], Lorad Selenia mam-

TABLE 4: Correlation of Results for 778 Examinations Obtained Using Fully Automated Volumetric Method With BI-RADS Category Assigned by Agreement of at Least Two Radiologists

BI-RADS Category	Fully Automated Volumetric Method[a] (Mean ± SD)		
	Fibroglandular Tissue Volume (cm^3)	Breast Tissue Volume (cm^3)	Breast Density (%)
1	33.7 ± 7.3	551.8 ± 76.4	6.1 ± 0.9
2	40.7 ± 13.3	560.3 ± 218.3	7.8 ± 2.3
3	50.7 ± 23.7	394.9 ± 179.8	14.1 ± 5.8
4	63.8 ± 35.4	265.4 ± 168.5	26.1 ± 5.2
p	< 0.001	< 0.001	< 0.001

[a]Volpara software, version 1.5.1, Mātakina Technology.

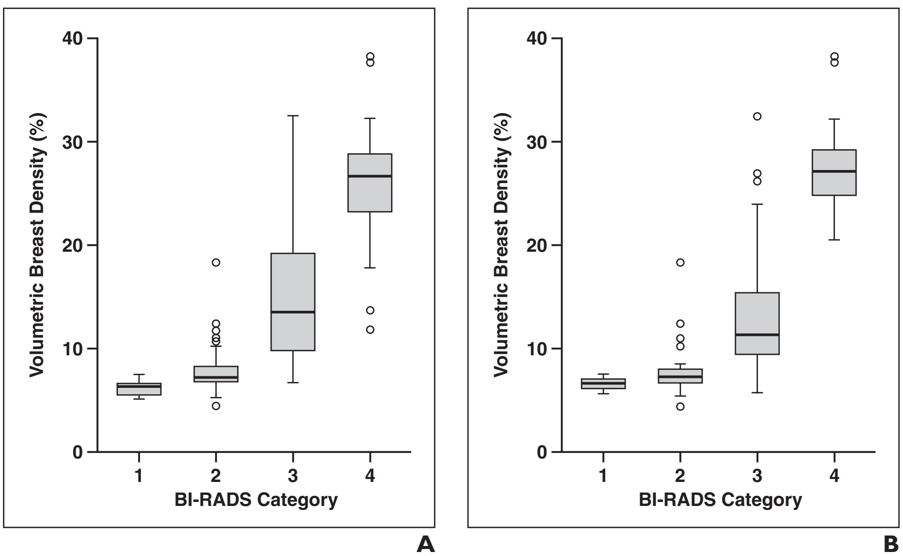

Fig. 2—Box-and-whisker plots show association between BI-RADS density categories assigned by radiologists and volumetric breast density assessed by fully automated volumetric method using Volpara software (version 1.5.1, Mātakina Technology). Whiskers show most extreme values within 1.5 interquartile ranges. Thick horizontal lines = median value, ○ = outliers.
A, Two radiologists in agreement about BI-RADS category and volumetric breast density by fully automated volumetric measurement.
B, Three radiologists in agreement about BI-RADS category and volumetric breast density by fully automated volumetric measurement.

mograms were considered less accurate for volumetric measurements than Senographe 2000D images because of a major difference between the two mammography systems: The Lorad Selenia compression paddle is designed to tilt during compression, whereas the compression paddle that is used in the Senographe 2000D is more rigid. The use of a flexible paddle results in variation in breast thickness measurements from the chest wall to the breast margin. However, our study showed no significant difference for volumetric breast density data between the two different systems.

The automated volumetric method could not produce reliable density measurements for women who are in special conditions that can increase breast density other than breast parenchyma, such as breast implantation or foreign-body injection. Although this proportion was small (1.5%), women who had breast implantation or foreign body injection had inaccurately high volumetric breast density values in our study. Breasts with lesions (mass or calcification) or breasts iatrogenically altered by cancer surgery also could affect automated measurements of volumetric density, but these associations have not yet been investigated. We assessed the effect of lesions (mass and calcification) and of breast cancer surgery on volumetric breast density by comparing breast density between the abnormal breast (mass or calcification) and normal contralateral breast in single patients, and no significant differences were shown in patients with breast lesions. However, volumetric breast density was significantly increased in breasts iatrogenically altered by cancer surgery. Therefore, if the automated volumetric method is to be used in a clinical setting, interpreters should take into account clinical history and various breast conditions.

As shown in Table 4, we found fibroglandular tissue volume and breast density were significantly increased and total breast tissue volume was significantly decreased in BI-RADS category 4. Considering that breast density can be affected by the volume of breast, the automated volumetric software may show increased density measurements for small breasts. However, the assignment of BI-RADS category 4 by observers is usually not influenced by breast size. Only a few studies have assessed the absolute dense breast area or volume [23, 26]. Further study will be needed for evaluation of absolute breast volume, fibroglandular volume, and percentage breast density as well as the association of breast size with breast density and breast cancer risk.

This study had some limitations. First, a reference standard to evaluate breast density does not exist. To address this issue, we assumed as a reference standard the BI-RADS density categories of at least two radiologists in agreement. Second, three radiologists in a single institution assigned BI-RADS density categories. It would be best to perform a larger study with more patients and radiologists from a variety of practice settings to validate our findings. Third, two different mammography units (Senographe 2000D and Lorad Selenia) were used. However, no significant differences in patient age, data from the automated software, and BI-RADS breast density category between the two different systems were noted (Table 1). Finally, the association of volumetric density with many factors affecting breast density was not investigated. A high-percentage breast density was strongly related to younger age, lower body mass index (BMI), nulliparity, late age at first delivery, and pre- and perimenopausal status [20, 27]. However, Lokate et al. [23] reported that high BMI was significantly related to low absolute dense area but at the same time was significantly related to high dense volume. A few other studies also reported that high BMI was related to larger dense volume on volumetric methods [28, 29]. The volumetric method could provide new insight into the relationship between BMI and other related factors and breast density.

In conclusion, our study showed good correlation of the fully automated volumetric method with radiologist-assigned BI-RADS density categories. The difficulties associated with the use of BI-RADS density categories, such as moderate to substantial interobserver agreement and broad quantitative range in the same category, may be avoided when mammographic density can be assessed using a fully automated volumetric method with digital mammography.

References

1. D'Orsi CJ, Mendelson EB, Ikeda DM, et al. *Breast Imaging Reporting and Data System: ACR BI-RADS—breast imaging atlas*. Reston, VA: American College of Radiology, 2003
2. Kerlikowske K, Grady D, Barclay J, et al. Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System. *J Natl Cancer Inst* 1998; 90:1801–1809
3. Berg WA, Campassi C, Langenberg P, Sexton MJ. Breast Imaging Reporting and Data System: inter- and intraobserver variability in feature analysis and final assessment. *AJR* 2000; 174:1769–1777
4. Martin KE, Helvie MA, Zhou C, et al. Mammographic density measured with quantitative computer-aided method: comparison with radiologists' estimates and BI-RADS categories. *Radiology* 2006; 240:656–665
5. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer* 1976; 37:2486–2492
6. Gram IT, Funkhouser E, Tabar L. The Tabar classification of mammographic parenchymal patterns. *Eur J Radiol* 1997; 24:131–136
7. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. *Phys Med Biol* 1994; 39:1629–1638
8. Sivaramakrishna R, Obuchowski NA, Chilcote WA, Powell KA. Automatic segmentation of mammographic density. *Acad Radiol* 2001; 8:250–256
9. Prevrhal S, Shepherd JA, Smith-Bindman R, Cummings SR, Kerlikowske K. Accuracy of mammographic breast density analysis: results of formal operator training. *Cancer Epidemiol Biomarkers Prev* 2002; 11:1389–1393
10. Highnam RP, Brady JM, Shepstone BJ. Mammographic image analysis. *Eur J Radiol* 1997; 24:20–32
11. Highnam R, Pan X, Warren R, Jeffreys M, Davey Smith G, Brady M. Breast composition measurements using retrospective standard mammogram form (SMF). *Phys Med Biol* 2006; 51:2695–2713
12. Jeffreys M, Warren R, Highnam R, Smith GD. Initial experiences of using an automated volumetric measure of breast density: the standard mammogram form. *Br J Radiol* 2006; 79:378–382
13. Tagliafico A, Tagliafico G, Astengo D, et al. Mammographic density estimation: one-to-one comparison of digital mammography and digital breast tomosynthesis using fully automated software. *Eur Radiol* 2012; 22:1265–1270
14. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159–174
15. Brisson J, Diorio C, Masse B. Wolfe's parenchymal pattern and percentage of the breast with mammographic densities: redundant or complementary classifications? *Cancer Epidemiol Biomarkers Prev* 2003; 12:728–732
16. Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). *Cancer Causes Control* 2000; 11:653–662
17. Boyd NF, Dite GS, Stone J, et al. Heritability of mammographic density, a risk factor for breast cancer. *N Engl J Med* 2002; 347:886–894
18. Ursin G, Ma H, Wu AH, et al. Mammographic density and breast cancer in three ethnic groups. *Cancer Epidemiol Biomarkers Prev* 2003; 12:332–338
19. Chen Z, Wu AH, Gauderman WJ, et al. Does mammographic density reflect ethnic differences in breast cancer incidence rates? *Am J Epidemiol* 2004; 159:140–147
20. Boyd NF, Guo H, Martin LJ, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med* 2007; 356:227–236
21. Harvey JA, Bovbjerg VE. Quantitative assessment of mammographic breast density: relationship with breast cancer risk. *Radiology* 2004; 230:29–41
22. Tagliafico A, Tagliafico G, Tosto S, et al. Mammographic density estimation: comparison among BI-RADS categories, a semi-automated software and a fully automated one. *Breast* 2009; 18:35–40
23. Lokate M, Kallenberg MG, Karssemeijer N, Van den Bosch MA, Peeters PH, Van Gils CH. Volumetric breast density from full-field digital mammograms and its association with breast cancer risk factors: a comparison with a threshold method. *Cancer Epidemiol Biomarkers Prev* 2010; 19:3096–3105
24. Tyson AH, Mawdsley GE, Yaffe MJ. Measurement of compressed breast thickness by optical stereoscopic photogrammetry. *Med Phys* 2009; 36:569–576
25. Mawdsley GE, Tyson AH, Peressotti CL, Jong RA, Yaffe MJ. Accurate estimation of compressed breast thickness in mammography. *Med Phys* 2009; 36:577–586
26. Dorgan JF, Klifa C, Shepherd JA, et al. Height, adiposity and body fat distribution and breast density in young women. *Breast Cancer Res* 2012; 14:R107
27. Stone J, Warren RM, Pinney E, Warwick J, Cuzick J. Determinants of percentage and area measures of mammographic density. *Am J Epidemiol* 2009; 170:1571–1578
28. Aitken Z, McCormack VA, Highnam RP, et al. Screen-film mammographic density and breast cancer risk: a comparison of the volumetric standard mammogram form and the interactive threshold measurement methods. *Cancer Epidemiol Biomarkers Prev* 2010; 19:418–428
29. Jeffreys M, Warren R, Highnam R, Davey Smith G. Breast cancer risk factors and a novel measure of volumetric breast density: cross-sectional study. *Br J Cancer* 2008; 98:210–216