Radiology

Breast Cancer Risk and Mammographic Density Assessed with Semiautomated and Fully Automated Methods and BI-RADS[1]

Abra M. Jeffers, PhD, MPhil
Weiva Sieh, PhD, MD
Jafi A. Lipson, MD
Joseph H. Rothstein, MS
Valerie McGuire, PhD
Alice S. Whittemore, PhD
Daniel L. Rubin, MD, MS

Purpose: To compare three metrics of breast density on full-field digital mammographic (FFDM) images as predictors of future breast cancer risk.

Materials and Methods: This institutional review board–approved study included 125 women with invasive breast cancer and 274 age- and race-matched control subjects who underwent screening FFDM during 2004–2013 and provided informed consent. The percentage of density and dense area were assessed semiautomatically with software (Cumulus 4.0; University of Toronto, Toronto, Canada), and volumetric percentage of density and dense volume were assessed automatically with software (Volpara; Volpara Solutions, Wellington, New Zealand). Clinical Breast Imaging Reporting and Data System (BI-RADS) classifications of breast density were extracted from mammography reports. Odds ratios and 95% confidence intervals (CIs) were estimated by using conditional logistic regression stratified according to age and race and adjusted for body mass index, parity, and menopausal status, and the area under the receiver operating characteristic curve (AUC) was computed.

Results: The adjusted odds ratios and 95% CIs for each standard deviation increment of the percentage of density, dense area, volumetric percentage of density, and dense volume were 1.61 (95% CI: 1.19, 2.19), 1.49 (95% CI: 1.15, 1.92), 1.54 (95% CI: 1.12, 2.10), and 1.41 (95% CI: 1.11, 1.80), respectively. Odds ratios for women with extremely dense breasts compared with those with scattered areas of fibroglandular density were 2.06 (95% CI: 0.85, 4.97) and 2.05 (95% CI: 0.90, 4.64) for BI-RADS and Volpara density classifications, respectively. Clinical BI-RADS was more accurate (AUC, 0.68; 95% CI: 0.63, 0.74) than Volpara (AUC, 0.64; 95% CI: 0.58, 0.70) and continuous measures of percentage of density (AUC, 0.66; 95% CI: 0.60, 0.72), dense area (AUC, 0.66; 95% CI: 0.60, 0.72), volumetric percentage of density (AUC, 0.64; 95% CI: 0.58, 0.70), and density volume (AUC, 0.65; 95% CI: 0.59, 0.71), although the AUC differences were not statistically significant.

Conclusion: Mammographic density on FFDM images was positively associated with breast cancer risk by using the computer assisted methods and BI-RADS. BI-RADS classification was as accurate as computer-assisted methods for discrimination of patients from control subjects.

[1] From the Departments of Management Science and Engineering (A.M.J.) and Medicine (Biomedical Informatics Research) (D.L.R.), Stanford University, Stanford, Calif; and Departments of Health Research and Policy (W.S., J.H.R., V.M., A.S.W.) and Radiology (J.A.L., D.L.R.), Stanford University School of Medicine, 1201 Welch Rd, Office P285, Stanford, CA 94305. Received September 17, 2015; revision requested October 29; revision received May 4, 2016; accepted May 24; final version accepted June 29. **Address correspondence to** D.L.R. (e-mail: *dlrubin@stanford.edu*).

Supported by National Institutes of Health (R01CA166827).

A.M.J. and W.S. contributed equally to this work.

Mammographic density refers to the opacity or white areas on a mammogram that represent parts of the breast containing large amounts of epithelial and stromal tissue relative to adipose tissue (1,2). The association between mammographic density and breast cancer risk was established in studies (3,4) in which investigators used film mammography. Estimated odds ratios ranged from 2.9 to 6.0, with the highest density categories compared with the lowest, and were approximately 1.5 for each standard deviation increment of mammographic density in studies of film mammograms (5).

Film mammography has been replaced with full-field digital mammography (FFDM) in recent years. Both film mammography and FFDM use x-rays to produce an image of the breast, which is saved directly on film or as digital data, respectively. In FFDM, there are two types of images: raw and processed. Raw images directly reflect x-ray absorption by the imaging detectors, which is processed by using the manufacturers' proprietary algorithms to improve the aesthetics and conspicuity of cancer on the processed images (6). If density is to be used for risk assessment at FFDM, the performance of different density estimation methods to predict breast cancer at FFDM must be evaluated.

Several measures of breast density have been developed. The American College of Radiology's Breast Imaging Reporting and Data System (BI-RADS) has four qualitative breast density categories: almost entirely fatty, scattered areas of fibroglandular density, heterogeneously dense, and extremely dense (7,8). A BI-RADS density classification is assigned by the radiologist on the basis of visual inspection of the image. The most common quantitative measurement software, Cumulus (University of Toronto, Toronto, Canada) (9,10), uses semiautomated, user-interactive thresholding of the image to estimate the percentage of breast area that is dense tissue. Cumulus is limited by intra- and interreader variability in establishing a threshold for segmenting dense tissue from surrounding fat, and the process of measuring density with this method is labor intensive and time consuming. Nonetheless, Cumulus is regarded as the reference standard for predicting breast cancer risk on the basis of quantitative assessment of digitized film mammograms (11).

The Cumulus method uses two-dimensional images of the breast; thus it is limited because breast density is three-dimensional and potentially variable in appearance on two-dimensional mammograms due to differences in compression and projection angle (12). A measure that allows estimation of the volume of fibroglandular tissue relative to the total breast is expected to enhance the association between breast cancer and breast density. Volpara (Volpara Solutions, Wellington, New Zealand) is a commercial product recently developed to provide automated volumetric measurements of breast density. Volpara also uses an algorithm approved by the U.S. Food and Drug Administration to produce density classifications similar to those in the fourth edition of the BI-RADS manual.

The relationship of mammographic density to breast cancer risk may vary depending on the measurement method. To our knowledge, few studies have been performed to compare the association between density estimated by using different methods for processing FFDM images and breast cancer risk (10,13,14). The objective of this study was to compare the association of density measurements obtained by using three different approaches (Cumulus, Volpara, and radiologist-assigned BI-RADS classification) with breast cancer risk.

Advances in Knowledge

- Mammographic density measured at full-field digital mammography was significantly associated with breast cancer risk when assessed by radiologists according to the Breast Imaging Reporting and Data System (BI-RADS) and by using semiautomated and fully automated computer-assisted methods.
- The area under the receiver operating characteristic curve for mammographic percentage of density was 0.68, 0.66, and 0.64 for BI-RADS, the semiautomated method, and the fully automated method, respectively; thus, BI-RADS was as accurate as computer-assisted methods for discrimination of patients with breast cancer from control subjects.

Implication for Patient Care

- BI-RADS should be considered an appropriate measure of mammographic density for estimation of patient risk of breast cancer and clinical decision making.

Materials and Methods

Study data were collected under a protocol approved by our institutional review board. Retrospective data were collected from patients who previously signed information release documents. Volpara Solutions provided the Volpara

Published online before print
10.1148/radiol.2016152062 **Content code:** BR

Radiology 2017; 282:348–355

Abbreviations:
AUC = area under the receiver operating characteristic curve
BI-RADS = Breast Imaging Reporting and Data System
BMI = body mass index
CI = confidence interval
FFDM = full-field digital mammography

Author contributions:
Guarantors of integrity of entire study, A.M.J., D.L.R.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, A.M.J., J.A.L.; clinical studies, A.M.J., J.A.L.; experimental studies, D.L.R.; statistical analysis, A.M.J., W.S., J.H.R., V.M., A.S.W.; and manuscript editing, all authors

Conflicts of interest are listed at the end of this article.

software. The authors had control of the data and the information submitted for publication.

Study Design and Population

We performed a matched case-control study to investigate the association of the percentage of density and absolute density with future risk of developing breast cancer on the basis of FFDM images assessed by using three density measurement approaches: Cumulus two-dimensional quantitative assessment, Volpara three-dimensional quantitative and qualitative assessment, and qualitative assessment by a radiologist according to BI-RADS criteria.

Patients comprised 125 women who underwent screening mammography at our institution between July 2003 and November 2012 and subsequently received a diagnosis of breast cancer. We selected the prediagnostic mammographic examination as the one at least 1 year before the date of diagnosis of breast cancer, when available, and closest to 1 year otherwise (median of 475 days from the prediagnostic mammographic examination to diagnosis). We used the image of the noncancerous breast contralateral to the affected breast to assess breast density, while avoiding the effect of the presence of cancer on the density measurement.

Control subjects comprised 274 women without a history of breast cancer who underwent screening mammography at our institution between 2004 and 2013. We ascertained the breast cancer–free status of control subjects by using the following inclusion criteria: *(a)* at least 10 years of follow-up for women aged 50 years or older at mammography and *(b)* three or more screening mammograms negative for cancer (BI-RADS assessment category 1 or 2) for women younger than 50 years at mammography. Exclusion criteria were: *(a)* diagnosis with breast cancer in our institutional cancer registry, *(b)* history of breast cancer in the medical record, *(c)* BI-RADS likelihood of malignancy assessment score of 3 or greater in mammography reports, or *(d)* breast cancer or breast implants noted in pathologic reports. Control subjects were frequency matched to patients according to 5-year age categories and race.

Acquisition of Images and Estimation of Density

All images were acquired at standard screening mammography. We selected the craniocaudal view of the noncancerous breast for patients and the left craniocaudal view for the matched control subjects. All mammograms were acquired with FFDM mammography units (Senograph Essential or Senograph 2000D; GE Medical Systems, Milwaukee, Wis). The patients were not stratified according to machine type, because this information (namely, the Digital Imaging and Communications in Medicine header for the "ManufacturerModelName" data element) was missing for the available images. These systems produce both raw and processed images, and the latter are used for clinical image display. The raw images have 14-bit dynamic range for each pixel (values ranging from 0 to 16383), whereas the processed images have 12-bit dynamic range (values ranging from 0 to 4095). The raw image pixel scale can be considered to represent the x-ray attenuation values, where adipose image regions appear bright (high pixel values) and fibroglandular regions appear dark (low pixel values). The processed images have a reversed intensity scale and a reduced dynamic range and resemble film mammograms.

To perform two-dimensional breast density assessment, the percentage of density measurements were estimated on the basis of the processed FFDM images in Digital Imaging and Communications in Medicine format by using Cumulus 6 software (9). To perform the Cumulus assessment, the operator adjusts a window and level to display the image optimally so that a threshold can be set to separate the dense tissue from the fatty tissue. The total breast area is determined by outlining the breast margins. The Cumulus percentage of density is calculated as the dense area divided by the total breast area. Processed GE images use only a small fraction of their 12-bit dynamic range, but are displayed in Cumulus at the full range (0–4095). To facilitate use of Cumulus on processed GE images, we applied the "softer" window boundaries by using the sigmoid transformation specified in the GE image metadata, and down-sampled images from a pixel size of 94 microns to 200 microns.

Study images for patients and their matched control subjects were read during the same session in random order, and a random subset of 10% of the study images (n = 48) were read a second time to assess intrareader reproducibility. A single reader (A.J., with 2 years of experience), who was blinded to whether the images were for patients or control subjects, performed all Cumulus measurements. The reader was trained by the providers of the Cumulus software (9). We previously found that noise reduction of processed FFDM images makes them appear more like film mammograms, and this can significantly improve reproducibility for readers with relatively little prior experience (15). Here, we applied a median filter with a radius of 2 pixels to assist the study reader (16). Reader proficiency was confirmed by means of a blinded test in which she attained excellent intrareader agreement, with a Pearson correlation coefficient of 0.90 (4,17), on the basis of 40 mammograms read twice in a different random order three days apart to reduce recall bias and an interreader correlation of 0.81, with the Cumulus instructor reading the same images. Reproducibility between Cumulus operators is generally lower than within the same operator (18). The proficiency of the study reader was independently confirmed in a blinded test of 74 study images that were read 5 months apart, in which she attained an intrareader correlation coefficient of 0.92.

To perform three-dimensional breast density assessment, we estimated Volpara volumetric percentage of density from the raw (unprocessed) images. We calculated percentage of density on the basis of raw mammograms as the total dense volume, normalized by the total breast volume, to give the percentage

of dense breast tissue as the volumetric percentage of density measure. In comparison with two-dimensional methods (BI-RADS density assessment and Cumulus), which do not account for the degree of thickness or compression of the breast, the Volpara three-dimensional method accounts for these factors in its estimation of volumetric density. The software also assigns classifications that are similar to the BI-RADS density categories 1–4, defined as less than 4.5%, 4.5%–7.49%, 7.5%–15.49%, and greater than or equal to 15.5% volumetric percentage of density (14). To obtain the clinical assessment of breast density, we recorded the BI-RADS category assigned by the radiologist in the mammographic report for each case.

Statistical Analysis

We used conditional logistic regression stratified by age (<40, 40–44, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, and 75–79 years) and race (white, black, Asian, other) to estimate the odds ratio and 95% confidence interval (CI) for breast cancer associated with the Cumulus percentage of density and dense area, Volpara volumetric percentage of density and dense volume, and BI-RADS and Volpara density categories, adjusting for menopausal status, parity, and body mass index (BMI). We selected variables by using backward stepwise conditional logistic regression with the Akaike information criterion (19).

Quantitative density measurements were categorized into quartiles on the basis of the density distribution in the control women. We applied square-root and cube-root transformations to Cumulus percentage of density and dense area, respectively, and log-transformations to Volpara volumetric percentage of density and dense volumes, to obtain normal distributions. Quantitative density measurements were also modeled as continuous variables in units of the standard deviation in the control women. To facilitate comparison with clinical BI-RADS assessments, we also categorized the quantitative density measurements by using the percentile cut points of the BI-RADS distribution in the control women. BMI (in kilograms per square meter) was categorized as less than 25, 25–29, and 30 plus. Menopausal status was categorized as premenopausal, perimenopausal, and postmenopausal. Parity was categorized as nulliparous and one, two, and three or more live births. We used multiple imputation techniques to handle missing values for parity and BMI (20). Assuming a multivariate normal distribution, each missing value was replaced with a set of plausible values by using the Markov chain Monte Carlo method. Each imputed dataset (n = 10) was then analyzed by using conditional logistic regression matched according to age and race and adjusted for menopausal status, parity, and BMI. The parameter estimates were then combined to produce a single risk estimate with a 95% CI.

A receiver operating characteristic curve for each density measure was created by plotting the true-positive rate versus the false-positive rate at varying density thresholds in models including age, race, menopausal status, parity, and BMI. The area under the receiver operating characteristic curve (AUC) is the probability that, in a randomly selected patient-control subject pair, the patient will have a higher assigned risk than does the control subject. The AUC is a measure of how well the metric discriminates between patients and control subjects. A value of 0.5 indicated that the model was no better than chance at making a prediction of membership in a group, and a value of 1.0 indicated that the model perfectly predicted membership in a group. We compared the AUCs for different density estimation methods by using the DeLong test (21). All analyses were performed with software (SAS version 9.4; SAS Institute, Cary, NC). A P value of less than or equal to .05 was considered to indicate a significant difference.

Results

Study Population

The study included 125 patients with invasive breast cancer and 274 control subjects (Table 1). Compared with control subjects, a larger proportion of the patients were postmenopausal and had a BMI greater than or equal to 30 kg/m^2, and a smaller proportion of the patients were nulliparous. Approximately 55% of patients and 39.8% of control subjects had heterogeneously dense or extremely dense breasts (clinical BI-RADS classification C or D). Cumulus area-based percentage of density measurements were substantially higher than were Volpara volumetric percentage of density measurements; however, the two percentage of density measures were highly correlated, with a Pearson correlation coefficient of 0.82. Cumulus dense area and Volpara dense volume were also correlated, with a Pearson correlation coefficient of 0.71. The agreement of clinical BI-RADS and Volpara density categorizations was fair, with a weighted κ statistic of 0.47.

Association of Breast Density with Breast Cancer Risk

Table 2 shows the associations of quantitative Cumulus and Volpara density measures and breast cancer risk, controlling for age, race, BMI, parity, and menopausal status. Compared with women in the second quartile, women with Cumulus percentage of density in the highest quartile had an odds ratio of 2.00 (95% CI: 1.01, 3.98; P = .047) and women in the highest quartile of Volpara volumetric percentage of density had an odds ratio of 1.71 (95% CI: 0.83, 3.53; P = .147). All four continuous measures were significantly associated with breast cancer risk, although associations with absolute density measures were generally weaker than those with percentage of density measures for both the Cumulus and Volpara methods. The odds ratio for each standard deviation increment was 1.61 (95% CI: 1.19, 2.19; P = .002) and 1.49 (95% CI: 1.15, 1.92; P = .002) for Cumulus percentage of density and dense area, and 1.54 (95% CI: 1.12, 2.10; P = .007) and 1.41 (95% CI: 1.11, 1.80; P = .005) for Volpara volumetric percentage of density and dense volume.

Table 1

Characteristics of Patients and Control Subjects

Characteristic	Patients (n = 125)	Control Subjects (n = 274)	P Value
Age at mammography (y)			.95
<50	52 (41.6)	113 (41.2)	
50+	73 (58.4)	161 (58.8)	
Menopausal status			.35
Premenopausal	37 (29.6)	94 (34.3)	
Postmenopausal	88 (70.4)	180 (65.7)	
Parity			.27
0	20 (21.5)	66 (26.7)	
1	18 (19.4)	42 (17.0)	
2	28 (30.1)	89 (36.0)	
3+	27 (29.0)	50 (20.3)	
Missing data	32	27	
BMI (kg/m^2)			.13
<25	58 (47.5)	135 (57.7)	
25–29	37 (30.4)	64 (27.3)	
30+	27 (22.1)	35 (15.0)	
Missing data	3	40	
BI-RADS classification			.007
A, Amost entirely fatty	10 (8.0)	53 (19.3)	
B, Scattered areas of fibroglandular density	46 (36.8)	112 (40.9)	
C, Heterogeneously dense	57 (45.6)	87 (31.8)	
D, Extremely dense	12 (9.6)	22 (8.0)	
Volpara density category			.43
1, Almost entirely fatty	20 (16.0)	54 (19.7)	
2, Scattered areas of fibroglandular density	19 (15.2)	54 (19.7)	
3, Heterogeneously dense	54 (43.2)	99 (36.1)	
4, Extremely dense	32 (25.6)	67 (24.5)	
Cumulus percentage of density quartile			.36
1, <12.12%	22 (17.6)	68 (24.8)	
2, 12.12%–25.19%	30 (24.0)	69 (25.2)	
3, 25.20%–36.34%	35 (28.0)	68 (25.2)	
4, ≥36.35%	38 (30.4)	69 (24.8)	
Mean*	28.1 ± 15.1	25.1 ± 14.9	.06
Cumulus dense area quartile			.13
1, <146.74 cm^2	19 (15.2)	68 (24.8)	
2, 146.74–269.64 cm^2	34 (27.2)	69 (25.2)	
3, 269.65–400.29 cm^2	31 (24.8)	68 (25.2)	
4, ≥400.30 cm^2	41 (32.8)	69 (24.8)	
Mean*	341.8 ± 208.5	288.1 ± 190.0	.01
Volpara volumetric percentage of density quartile			.79
1, <5.20%	27 (21.6)	69 (25.2)	
2, 5.20%–8.83%	29 (23.2)	68 (24.8)	
3, 8.84%–14.81%	36 (28.8)	68 (24.8)	
4, ≥14.82%	33 (26.4)	69 (25.2)	
Mean*	12.3 ± 8.4	11.2 ± 7.9	.21
Volpara dense volume quartile			.41
1, <34.21 cm^3	24 (19.2)	68 (24.8)	
2, 34.21–49.99 cm^3	33 (26.4)	69 (25.2)	
3, 50.00–73.65 cm^3	28 (22.4)	68 (25.2)	
4, ≥73.66 cm^3	40 (32.0)	69 (24.8)	
Mean*	67.4 ± 42.3	57.0 ± 32.6	.02

Note.—Unless otherwise indicated, data are number of patients, with percentage in parentheses.

*Data are means ± standard deviation.

Table 3 shows the association of clinical BI-RADS and fully automated Volpara density classifications with breast cancer risk, controlling for age, race, BMI, parity, and menopausal status. Compared with women with areas of scattered fibroglandular density (BI-RADS B), women with fatty breasts had reduced risk (odds ratio, 0.38; 95% CI: 0.17, 0.84; P = .017) and women with heterogeneously dense (odds ratio, 2.35; 95% CI: 1.34, 4.12; P = .003) or extremely dense (odds ratio, 2.06; 95% CI: 0.85, 4.97; P = .107) breasts had elevated risk. Associations were generally weaker for Volpara classifications than for clinical BI-RADS classifications, although a statistically significant trend of increasing breast cancer risk with increasing density categories was observed for both methods.

To facilitate comparisons between clinical BI-RADS and quantitative density measurements, we categorized Cumulus and Volpara measurements by using the percentile cut points of the clinical BI-RADS distribution in the control women (Table 4). Compared with women in the second BI-RADS-like category, women in the highest category had more than twofold increased risks of breast cancer for Cumulus percentage of density and dense area, and Volpara volumetric percentage of density and dense volume measurements. The positive trend toward increasing risks with increasing density was most statistically significant for clinical BI-RADS assessments.

Discrimination of Patients and Control Subjects

The Figure shows the receiver operating characteristic curves for clinical BI-RADS, Cumulus, and Volpara density measurements. Clinical BI-RADS assessment was slightly better for discrimination of patients from control subjects (AUC, 0.68; 95% CI: 0.63, 0.74) than was Volpara density classification (AUC, 0.64; 95% CI: 0.58, 0.70) or quantitative measures of Cumulus percentage of density (AUC, 0.66; 95% CI: 0.60, 0.72) and dense area (AUC, 0.66; 95% CI: 0.60, 0.72), and Volpara volumetric percentage

Table 2

Breast Cancer Risk Associated with Quantitative Mammographic Density Measured with Cumulus and Volpara Methods

	Cumulus		Volpara	
Density Quartile or Variable	Percentage of Density	Dense Area	Volumetric Percentage of Density	Dense Volume
Quartile 1, ≤25% dense	0.64 (0.33, 1.27)	0.47 (0.23, 0.97)*	0.75 (0.37, 1.51)	0.67 (0.35, 1.28)
Quartile 2, 26%–50% dense	1.00	1.00	1.00	1.00
Quartile 3, 51%–75% dense	1.48 (0.78, 2.81)	0.89 (0.48, 1.66)	1.40 (0.74, 2.66)	0.86 (0.46, 1.60)
Quartile 4, >75% dense	2.00 (1.01, 3.98)*	1.33 (0.73, 2.41)	1.71 (0.83, 3.53)	1.35 (0.73, 2.51)
P value for trend	.004	.01	.05	.08
Each standard deviation increment	1.61 (1.19, 2.19)†	1.49 (1.15, 1.92)†	1.54 (1.12, 2.10)*	1.41 (1.11, 1.80)*

Note.—Unless otherwise indicated, data are odds ratios, with 95% CIs in parentheses. Odds ratios and CIs were estimated with conditional logistic regression, with age and race matched, and were adjusted for menopausal status, parity, and BMI.

* $P < .05$.

† $P < .005$.

Table 3

Breast Cancer Risk Associated with BI-RADS and Volpara Density Categories

Density Category	BI-RADS	Volpara
Almost entirely fatty	0.38 (0.17, 0.84)*	0.77 (0.35, 1.67)
Scattered areas of fibroglandular density	1.00	1.00
Heterogeneously dense	2.35 (1.34, 4.12)†	1.67 (0.85, 3.26)
Extremely dense	2.06 (0.85, 4.97)	2.05 (0.90, 4.64)
P value for trend	<.001	.02

Note.—Unless otherwise indicated, data are odds ratios, with 95% CIs in parentheses. Odds ratios and CIs were estimated with conditional logistic regression, with age and race matched, and were adjusted for menopausal status, parity, and BMI.

* $P < .05$.

† $P < .005$.

of density (AUC, 0.64; 95% CI: 0.58, 0.70) and dense volume (AUC, 0.65; 95% CI: 0.59, 0.71). However, the differences between the AUCs for clinical BI-RADS and other density measures were not statistically significant.

Discussion

In this case-control study of breast density assessed with Cumulus, Volpara, and BI-RADS from processed FFDM images, we found that all density measures were positively associated with breast cancer risk. Clinical assessment with BI-RADS allowed the best discrimination of patients from control subjects, followed by Cumulus and Volpara measurements, although the AUC differences were small and not statistically significant. In a recent study, Brandt et al (14) also reported that clinical assessment with BI-RADS based on FFDM images was slightly better for discrimination than was assessment with Volpara BI-RADS categories. Eng et al (10) found that clinical BI-RADS, Cumulus percentage of density, and Volpara volumetric percentage of density measurements based on FFDM images were all positively associated with breast cancer risk; although in their study, the AUCs were similar for Volpara and Cumulus percentage of density measures and were not reported for BI-RADS assessments.

Clinical BI-RADS scores are qualitative determinations of breast density made by radiologists. These assessments may be influenced by imaging characteristics introduced into the processed mammograms that are not present on the raw images. These differences could partially explain the fair agreement between clinical BI-RADS and Volpara density classifications and the slightly weaker associations with breast cancer risk when the fully automated Volpara measures are used.

The AUC is well known to be a conservative measure for detecting subtle differences in discriminatory ability (10,22–24), and it is unusual to see a substantial increase in the AUC, even with the addition of a very strong risk factor (25). In addition, it is difficult to determine the reference standard for breast density estimation on FFDM images. However, our results are interesting in that, if visual assessment is the standard, as is now legislated in more than half of all states in the United States, then Cumulus and Volpara do not appear to provide much additional value for prediction of risk beyond that of routine clinical BI-RADS assessment. Additional studies would be helpful to determine whether there may be small incremental benefits of current computerized methods. Beyond differences in risk prediction, automated methods may provide benefits of reducing interobserver variability and improving reproducibility of breast density assessments over time. Future development of automated methods to quantify breast density and other risk features could improve risk prediction with FFDM images.

A main strength of this study was the availability of both the raw and

Table 4

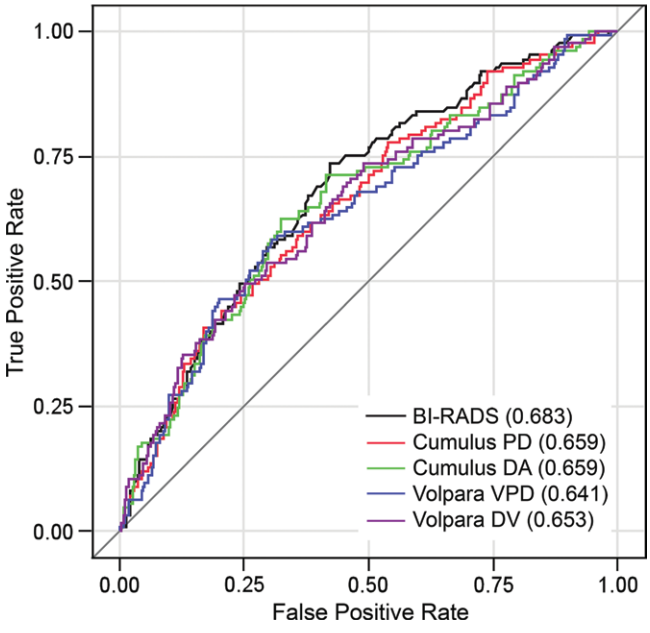
Breast Cancer Risk Associated with Cumulus fand Volpara Density Categories Defined by Using the Percentile Cutpoints of Radiologist BI-RADS Assessments in Control Subjects

		Cumulus		Volpara	
Density Category	BI-RADS	Percentage of Density	Dense Area	Volumetric Density Percentage	Dense Volume
Almost entirely fatty, ≤19%	0.38 (0.17, 0.84)*	0.68 (0.34, 1.35)	0.40 (0.18, 0.86)*	0.70 (0.36, 1.38)	0.72 (0.38, 1.38)
Scattered areas of fibroglandular density, 20%–60%	1.00	1.00	1.00	1.00	1.00
Homogeneously dense, 61%–92%	2.35 (1.34, 4.12)†	1.83 (1.06, 3.15)*	1.06 (0.63, 1.78)	1.75 (0.99, 3.11)	0.95 (0.54, 1.66)
Extremely dense, >92%	2.06 (0.85, 4.97)	2.09 (0.85, 5.11)	2.12 (0.99, 4.55)	2.76 (1.14, 6.69)*	2.18 (1.08, 4.39)*
P value for trend	<.001	.006	.003	.005	.03

Note.—Unless otherwise indicated, data are odds ratios, with 95% CIs in parentheses. Odds ratios and CIs were estimated with conditional logistic regression, with age and race matched, and were adjusted for menopausal status, parity, and BMI.

* $P < .05$.

† $P < .005$.



Graph shows AUCs for mammographic density measures. *PD* = percentage of density, *DA* = dense area, *VPD* = volumetric percentage of density, *DV* = dense volume.

processed mammographic images and corresponding radiologic reports, which allowed us to compare Cumulus, Volpara, and BI-RADS density assessments on all study images. Most institutions store only the processed images, and not the raw images, which prohibits retrospective Volpara assessments. A second important strength was the use of screening mammograms and images of the unaffected breast before the diagnosis of cancer in the contralateral breast to avoid the influence of the presence of cancer on the density assessments. Finally, this study included the collection of covariate data that were electronically captured at the time of mammographic screening from the electronic medical record.

There were several limitations to our study. First, the available sample size limited the ability to detect subtle differences in discrimination among the density assessment methods. Second, clinical BI-RADS density assessment was made by a single reader, as is the standard clinical practice in the United States. The purpose of our study was to compare routine clinical BI-RADS assessments, including variability among radiologists, with semiautomated and fully automated methods; thus, we did not require multiple radiologists to assess each image, although this could have improved the accuracy of the BI-RADS assessments. Finally, the Cumulus assessments were performed by a single reader, which has the advantage of eliminating interreader variability. The standard of practice for using Cumulus software is to require the reader to undergo specialized training and attain high levels of intrareader reproducibility with test images before reading the study images. The extensive training and time required to perform Cumulus measurements made it impractical to have more than one Cumulus reader for this study, although we acknowledge that having multiple readers could have strengthened the results.

Mammographic density assessed on the basis of FFDM images with

Cumulus, Volpara, and BI-RADS was significantly associated with breast cancer risk. Radiologist-assigned BI-RADS density categorization allowed discrimination of patients with breast cancer from control subjects at least as well as did computer-assisted methods.

Acknowledgment: We thank Volpara Solutions, Wellington, New Zealand, for the use of Volpara.

Disclosures of Conflicts of Interest: A.M.J. disclosed no relevant relationships. **W.S.** disclosed no relevant relationships. **J.A.L.** disclosed no relevant relationships. **J.H.R.** disclosed no relevant relationships. **V.M.** disclosed no relevant relationships. **A.S.W.** disclosed no relevant relationships. **D.L.R.** disclosed no relevant relationships.

References

1. Li T, Sun L, Miller N, et al. The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. Cancer Epidemiol Biomarkers Prev 2005;14(2):343–349.
2. Johns PC, Yaffe MJ. X-ray characterisation of normal and neoplastic breast tissues. Phys Med Biol 1987;32(6):675–695.
3. Boyd NF, Martin LJ, Bronskill M, Yaffe MJ, Duric N, Minkin S. Breast tissue composition and susceptibility to breast cancer. J Natl Cancer Inst 2010;102(16):1224–1237.
4. Vachon CM, van Gils CH, Sellers TA, et al. Mammographic density, breast cancer risk and risk prediction. Breast Cancer Res 2007;9(6):217.
5. Pettersson A, Graff RE, Ursin G, et al. Mammographic density phenotypes and risk of breast cancer: a meta-analysis. J Natl Cancer Inst 2014;106(5):dju078.
6. Pisano ED, Zuley M, Baum JK, Marques HS. Issues to consider in converting to digital mammography. Radiol Clin North Am 2007;45(5):813–830, vi.
7. Nicholson BT, LoRusso AP, Smolkin M, Bovbjerg VE, Petroni GR, Harvey JA. Accuracy of assigned BI-RADS breast density category definitions. Acad Radiol 2006;13(9):1143–1149.
8. Boyd NF, Martin LJ, Yaffe MJ, Minkin S. Mammographic density and breast cancer risk: current understanding and future prospects. Breast Cancer Res 2011;13(6):223.
9. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. Phys Med Biol 1994;39(10):1629–1638.
10. Eng A, Gallant Z, Shepherd J, et al. Digital mammographic density and breast cancer risk: a case-control study of six alternative density assessment methods. Breast Cancer Res 2014;16(5):439.
11. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006;15(6):1159–1169.
12. Yaffe MJ. Mammographic density. Measurement of mammographic density. Breast Cancer Res 2008;10(3):209.
13. Fowler EE, Vachon CM, Scott CG, Sellers TA, Heine JJ. Automated percentage of breast density measurements for full-field digital mammography applications. Acad Radiol 2014;21(8):958–970.
14. Brandt KR, Scott CG, Ma L, et al. Comparison of clinical and automated breast density measurements: implications for risk prediction and supplemental screening. Radiology 2016;279(3):710–719.
15. Habel LA, Lipson JA, Achacoso N, et al. Case-control study of mammographic density and breast cancer risk using processed digital mammograms. Breast Cancer Res 2016;18(1):53.
16. Gonzalez RC, Woods RE. Digital image processing. 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall, 2008.
17. Boyd NF, Rommens JM, Vogt K, et al. Mammographic breast density as an intermediate phenotype for breast cancer. Lancet Oncol 2005;6(10):798–808.
18. Keller BM, Nathan DL, Gavenonis SC, Chen J, Conant EF, Kontos D. Reader variability in breast density estimation from full-field digital mammograms: the effect of image postprocessing on relative and absolute measures. Acad Radiol 2013;20(5):560–568.
19. Draper N, Smith H. Applied regression analysis. 2nd ed. New York, NY: Wiley, 1976.
20. Berglund P, Heeringa S. SAS Institute. Multiple imputation of missing data using SAS. Cary, NC: SAS Institute, 2014.
21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44(3):837–845.
22. Li J, Szekely L, Eriksson L, et al. High-throughput mammographic-density measurement: a tool for risk prediction of breast cancer. Breast Cancer Res 2012;14(4):R114.
23. Sovio U, Li J, Aitken Z, et al. Comparison of fully and semi-automated area-based methods for measuring mammographic density and predicting breast cancer risk. Br J Cancer 2014;110(7):1908–1916.
24. Darabi H, Czene K, Zhao W, Liu J, Hall P, Humphreys K. Breast cancer risk prediction and individualised screening based on common genetic variation and breast density measurement. Breast Cancer Res 2012;14(1):R25.
25. Vickers AJ, Cronin AM, Begg CB. One statistical test is sufficient for assessing new predictive markers. BMC Med Res Methodol 2011;11:13.