

Evaluating the Effectiveness of Using Standard Mammogram Form to Predict Breast Cancer Risk: Case-Control Study

Jane Ding,¹ Ruth Warren,¹ Iqbal Warsi,¹ Nick Day,² Deborah Thompson,² Michael Brady,³ Christopher Tromans,³ Ralph Highnam,⁴ and Douglas Easton²

¹Department of Radiology, University of Cambridge, Addenbrooke's Hospital; ²Strangeways Research Laboratory, Cambridge, United Kingdom; ³Department of Engineering Science, University of Oxford, Oxford, United Kingdom; and ⁴Highnam Associates Limited, Wellington, New Zealand

Abstract

Breast density is a well-known breast cancer risk factor. Most current methods of measuring breast density are area based and subjective. Standard mammogram form (SMF) is a computer program using a volumetric approach to estimate the percent density in the breast. The aim of this study is to evaluate the current implementation of SMF as a predictor of breast cancer risk by comparing it with other widely used density measurement methods. The case-control study comprised 634 cancers with 1,880 age-matched controls combined from the Cambridge and Norwich Breast Screening Programs. Data collection involved assessing the films based both on Wolfe's parenchymal patterns and on visual estimation of percent density and then digitizing the films for computer analysis (interactive threshold technique and SMF). Logistic regression was used to produce odds ratios associated with increasing categories of breast

density. Density measures from all four methods were strongly associated with breast cancer risk in the overall population. The stepwise rises in risk associated with increasing density as measured by the threshold method were 1.37 [95% confidence interval (95% CI), 1.03-1.82], 1.80 (95% CI, 1.36-2.37), and 2.45 (95% CI, 1.86-3.23). For each increasing quartile of SMF density measures, the risks were 1.11 (95% CI, 0.85-1.46), 1.31 (95% CI, 1.00-1.71), and 1.92 (95% CI, 1.47-2.51). After the model was adjusted for SMF results, the threshold readings maintained the same strong stepwise increase in density-risk relationship. On the contrary, once the model was adjusted for threshold readings, SMF outcome was no longer related to cancer risk. The available implementation of SMF is not a better cancer risk predictor compared with the thresholding method. (Cancer Epidemiol Biomarkers Prev 2008;17(5):1074-81)

Introduction

Mammographic breast density refers to the radiographically dense regions on the mammogram, primarily associated with fibroglandular tissue. Breast density is one of the strongest risk factors for breast cancer. Women with high breast density have a 4- to 6-fold increased risk of developing breast cancer compared with those with low density (1-8). Breast density is currently undergoing evaluation as a biomarker for breast cancer, has great potential in its use as an intermediate end point in breast cancer research, and may provide important insights in the study of breast cancer etiology and disease pathways.

The first method of measuring breast density was proposed by Wolfe in 1976 and involves categorically classifying density (9). Later methods in the same vein include Tabar's five-point grading system, Breast Imaging Reporting and Data System, and the six-category classification method, which involves visually quantifying the extent of density covering the breast. Several

teams also visually estimated the percentage of dense breast tissue in 5% density increments, in other words, on a 20- or 21-category scale (3, 10). As precision became important in the quantification of breast density, Byng and colleagues (11) in Canada developed a computer-assisted thresholding technique for measuring density. Today, this and similar techniques are the most widely used method for calculating density.

The concerns with all of the current density measurement methods include reader subjectivity, labor intensiveness, differences in mammographic techniques, and lack of explicit consideration for the third dimension of the breast. The standard mammogram form (SMF) is a recently developed software that theoretically overcomes the previous addressed problems by estimating the amount of breast density volumetrically through a fully automated process.

SMF density measures have shown reasonable correlation with the visual six-category density classification method (12). However, it does have lower left-right reliability compared with the thresholding technique (13). Validation of a breast density measurement method should involve evaluating its ability to predict breast cancer risk in a case-control setting. In this study, we assess the association between breast density with cancer risk by comparing SMF with the conventional visual classification and the semiautomated density method.

Received 9/28/07; revised 2/12/08; accepted 2/25/08.

Grant support: EC grant for e-Science for the MammoGrid project. J. Ding was funded by the Overseas Research Scholarship and Cambridge Commonwealth Trust Scholarship throughout her Ph.D.

Requests for reprints: Jane Ding, Department of Radiology, University of Cambridge, Addenbrooke's Hospital, Box 218, Hills Road, Cambridge CB2 0QQ, United Kingdom. Phone: 44-1223-256119; Fax: 44-1223-330915. E-mail: jjd32@cam.ac.uk

Copyright © 2008 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-07-2634

Materials and Methods

Population Characteristics. Study subjects included women who were diagnosed with various stages of invasive breast cancer and attended screenings at the UK National Health Service Breast Cancer Screening Program in Cambridge between November 1995 and August 2003, and in Norwich and Norfolk between March 1998 and March 2004. A total of 1,271 cancers in Norwich and 221 cancers in Cambridge were identified as potential cancers. We excluded 858 women whose diagnosis was sole *in situ* cancer, whose age was outside screening range, whose films were not found, or whose films did not have the requisite film acquisition data, leaving a total of 634 cancer cases in the final analyses. Patients were divided into screen-detected patients, whose cancer was diagnosed at any screening round, and interval cancer patients, who were diagnosed symptomatically between screening rounds, after a negative screen.

In the United Kingdom, women ages 50 to 70 y receive invitations to attend routine breast screening every 3 y. Women ages >70 y are entitled to request screening if they so choose. Subjects in this study were ages between 50 and 75 y.

Controls were selected by searching the screening database from the National Health Service Regional Breast Screening Program for Cambridge and Norfolk and Norwich separately. For the Cambridge study, one control was selected for each cancer case. Approximately three controls were chosen for each cancer case in the Norwich project. The methods for control selection were exactly the same at both screening centers. For each screen-detected case, we selected controls who had also attended screening and been diagnosed as normal at that screen. For each interval case, we used the mammograms taken at the screening round before the date of the diagnostic mammogram as the basis for matching of controls. Controls were matched to the patients by screening center, date of birth (within 6 mo), and date of screening (within 3 mo). Patient characteristics at the Cambridge and Norwich regions are outlined in Table 1.

This study was approved by the Norfolk Local Research Ethics Committee. This was a medical records study; therefore, direct consent from the patients was not required.

Density Measurements. For cancer patients, we used mammograms from the contralateral or unaffected breast, and for the controls, we chose the matching side of the breast. We used only the mediolateral oblique view in the analyses as this single view was available consistently on all subjects (14). A total of 1,297 subjects

received imaging for the craniocaudal view, and the density results from these images were compared using the Spearman's rank correlation to determine the consistency between the two views. Density assessments were done by an expert radiologist (R.W.), who is experienced in evaluating parenchymal patterns and visually estimating percent density, and received training on use of the interactive threshold program from its originators. The reader did not have any knowledge of whether the image belonged to a cancer case or a control. All repeated readings were randomly selected and done by R. Warren within 3 wk of the first readings. R. Warren was blinded to previous assignments during these repeated readings.

Wolfe's Parenchymal Patterns. Mammographic density was visually assessed based on Wolfe's parenchymal patterns. An N1 pattern involved parenchymal that is composed primarily of fat with no visible ducts. Parenchyma in P1 pattern contained prominent ducts extending to up to 25% of the breast, and in P2, pattern involved ducts in >25% of breast area. DY pattern is related to severe involvement of the parenchyma with extensive amount of density. Intraobserver agreement was analyzed using the weighted κ test based on repeated reading of 227 films.

Visual Percent Density. The percentage of mammographic density covering the breast was estimated visually for every 5% density increment. The visual percent density scores were there for 0%, 5%, 10%, 15%, and so forth up to 95% and 100%. This 21-point categorical scoring system resembles a continuous outcome. A total of 223 images were used to determine the intraobserver agreement, which was analyzed by Spearman's rank correlation.

Interactive Threshold. Breast density was measured using the interactive thresholding software, which calculates the percent area of breast covered in dense regions. Original mammograms were scanned using the Array 2905 DICOM ScanPro Plus Laser Film Digitizer version 1.3E (Array Corp. USA) at absorbance of 4.7. Interactive thresholding measurements were done using the Cumulus 3.0 program designed by Yaffe and Boyd, the originators of the technique (15). This computer-assisted method uses the gray scale in the digitized images to determine breast density. The assessor first adjusts the threshold to mark the edge of the breast and to mask the pectoral muscle. A second threshold is set to outline the regions of density within the breast. The program computes the number of pixels in dense area and total breast area, from which percent of density is calculated. A selection of 150 images was reanalyzed and used to determine the intraobserver agreement, as calculated by Spearman's rank correlation.

Standard Mammogram Form. Volumetric density measurements were obtained using the software GenerateSMF version 2.2 β from Siemens Molecular Imaging. SMF provides a representation of the amount of nonfat tissue at each location in a mammogram, a representation that has been estimated by an evolving series of computer programs. If the separation between the mammography compression plates is known, then the SMF representation potentially provides a volume-based estimate of the amount of dense tissue in a breast. The

Table 1. Attributes of combined case-control studies

	Cambridge	Norwich
Cancers/controls (<i>n</i>)	140/140	494/1,740
Screen-detected	94/94	334/1,161
cancers/controls (<i>n</i>)		
Interval cancers/controls (<i>n</i>)	46/46	160/579
Mean age (y)	57.6	57.4
Age range (y)	50-75	50-75
Period of selection	1995-2003	1998-2004

SMF software available in this project requires knowledge of several film imaging variables, including tube voltage and film exposure time, to estimate the thickness of the compressed breast; this is a crucial step toward building a 2.5 dimensional model of the breast.⁵ Fortunately, it has recently been shown that SMF can equally be estimated without compromising much accuracy (16) for the majority of films that lack such imaging variable information. As the name SMF suggests, the SMF program compensates for variations in X-ray imaging techniques and standardizes image scanned on different film digitizers. The technical details of the SMF software have been described elsewhere (16, 17).

The Array-scanned images were converted into DICOM files compatible for use with SMF. X-ray calibration data, including tube voltage, film exposure time, anode target material, and filter material, were extracted from individual film labels and entered into the data entry section of the scanning software (software created by Christopher Tromans). All of the Norwich mammograms contained full set of calibration data as we excluded earlier films that did not have this information. The Cambridge mammograms were taken on two different types of mammography machines, one of which did not provide any calibration information on the film labels. Consequently, a total of 124 of 2,514 films (4.9%) did not have calibration data but were still included in the final analyses. The SMF software provides an estimate for breast thickness for each film and in doing so compensates for erroneous or lack of calibration data (16). The default calibration settings for films with unknown variables include tube voltage at 28 kVp, molybdenum/molybdenum anode target and filter combination, and estimation of exposure time from the scanned image.

The SMF program outputs the percent volume of density including and excluding the breast edge. Because the edge of the breast has inconsistent thickness compared with the rest of the breast and contains predominantly fatty tissue, we examined in this study only the values excluding the breast edge. Inevitably, as a result of this exclusion, the final percent volume of the breast as estimated by GenerateSMF is expected to be slightly understated.

Repeatability of SMF density measures was carried out on the full set of data (2,514 films) and a 100% correlation was achieved. This was expected as the software is fully automated.

The SMF software standardizes for different digitizers, given that the absorbance of scanned images is known and is either 4.0 or 4.7 (18). To test the digitizer effect, we compared SMF results on images scanned on an old hand-feed Lumysis digitizer (images scanned at absorbance 4.0) and on a modern Array machine (images scanned at absorbance 4.7; ref. 18).

Statistical Analyses. The association between mammographic density and breast cancer risk was evaluated

using logistic regression to compute odds ratios (OR) with 95% confidence intervals (95% CI). Age (three categories: 50-54, 55-59, and 60+), screening center (Cambridge versus Norwich), and mode of detection (screen-detected cancers versus interval cancers) were included into the multivariate analyses as confounders. We were unable to obtain information on body mass index and reproductive history as they were not recorded in the National Health Service Breast Screening Program and, therefore, could not include these as confounders into the multivariate analyses. Density values, as measured using the visual method, thresholding technique, and SMF, were divided into quartiles of combined cancers and controls to be compared with Wolfe's four-point grading system. For all methods, the three higher density groups were compared individually to the lowest baseline group. The fit of the models was compared using the likelihood ratio test statistics. The analyses were repeated assuming a linear trend in the relationship between density and risk (the trend test). The same analyses were carried out in separate screen-detected and interval cancer populations. Finally, we considered the model including both the SMF results and the threshold readings to determine whether each was an independent determinant of risk, considering breast density both as a continuous and a category variable. We conducted all statistical analyses using Stata 8.0 (Stata Corp.).

Quality Control of Density Readings. A total of 1,297 subjects received imaging for the craniocaudal view, which was additionally assessed by the reader. The SMF mammographic mediolateral oblique and craniocaudal view Spearman's rank correlation was 0.746 for percent volume of density. View correlations for the thresholding method were 0.962 for total breast area, 0.897 for total dense area, and 0.907 for percent area of density. A random selection of 150 images was used to assess reader repeatability for the thresholding method: the Spearman's rank correlations were 0.995 for total breast area, 0.930 for total dense area, and 0.933 for percent area of density. The Spearman correlation for visual estimates of percent density on 223 images was 0.957. The intraobserver agreement of the reader for parenchymal pattern assessment on 227 images was excellent (weighted $\kappa = 0.820$).

Results

A total of 2,514 women were included in this study, 634 of which were breast cancer patients and 1,880 of which were controls. Table 1 shows the basic characteristics of the Cambridge and Norwich study populations.

Table 2 summarizes the ORs associated with increasing categories of density in the overall population. In general, mean percent densities as measured by interactive threshold and SMF were higher in the cancer population than the control population, but the difference was more pronounced in the threshold method (Table 2). Strong density-risk trends existed in all four methods, as was shown by the *P* values for trend being <0.001. However, the stepwise increase in risk associated with higher density categories was greater in

⁵ It is called 2.5D because, at a particular image location *x*, the SMF value may be estimated to be 3.7 cm of nonfat tissue, but if the compressed breast thickness is 6 cm, the program cannot say which part of the 6-cm column of tissue corresponding to *x* is fat (2.3 cm) and which is nonfat.

Table 2. Association between breast density and risk in the overall, screen-detected, and interval cancer populations

	Cases (n)	Controls (n)	OR*	P	95% CI
Overall population					
Wolfe's patterns					
N1	116	465	1 (reference)	—	—
P1	137	455	1.16	0.316	0.87-1.54
P2	296	780	1.48	0.002	1.15-1.89
DY	85	180	1.79	0.001	1.27-2.52
P for trend			1.22 [†]	<0.001	
Visual percent density					
0-15%	180	674	1 (reference)	—	—
20-35%	100	353	1.00	0.979	0.76-1.33
40-65%	160	464	1.24	0.089	0.97-1.59
70-90%	194	389	1.81	<0.001	1.41-2.32
P for trend			1.21 [†]	<0.001	
Mean density, % (SD)	42.8 (28.6)	35.9 (27.9)			
Interactive threshold					
0.00-9.41%	112	517	1 (reference)	—	—
9.43-22.92%	144	484	1.37	0.029	1.03-1.82
22.94-37.03%	171	458	1.80	<0.001	1.36-2.37
37.04-84.74%	207	421	2.45	<0.001	1.86-3.23
P for trend			1.34 [†]	<0.001	
Mean density, % (SD)	29.3 (19.2)	23.8 (17.9)			
SMF					
11.5-19.1%	137	500	1 (reference)	—	—
19.2-22.7%	142	492	1.11	0.448	0.85-1.46
22.8-28.2%	155	461	1.31	0.054	1.00-1.71
28.3-59.3%	200	427	1.92	<0.001	1.47-2.51
P for trend			1.24 [†]	<0.001	
Mean density, % (SD)	25.6 (8.0)	24.0 (7.1)			
Screen-detected cancer population					
Wolfe's patterns					
N1	92	295	1 (reference)	—	—
P1	103	304	1.04	0.828	0.75-1.44
P2	186	528	1.08	0.614	0.80-1.45
DY	47	128	1.14	0.548	0.75-1.73
P for trend			1.04 [†]	0.502	
Visual percent density					
0-15%	142	435	1 (reference)	—	—
20-30%	61	209	0.81	0.237	0.57-1.15
35-60%	105	311	0.97	0.859	0.72-1.31
65-90%	120	300	1.18	0.277	0.88-1.58
P for trend			1.06 [†]	0.266	
Mean density, % (SD)	38.5 (28.0)	36.6 (27.6)			
Interactive threshold					
0.00-9.34%	91	330	1 (reference)	—	—
9.35-22.72%	107	314	1.23	0.222	0.88-1.70
22.73-36.13%	112	309	1.34	0.079	0.97-1.86
36.25-84.74%	118	302	1.56	0.007	1.13-2.17
P for trend			1.15 [†]	0.007	
Mean density, % (SD)	26.2 (18.3)	24.3 (18.1)			
SMF					
11.5-19.0%	111	323	1 (reference)	—	—
19.1-22.5%	94	315	0.90	0.533	0.65-1.25
22.6-28.1%	105	316	1.03	0.865	0.75-1.41
28.2-59.3%	118	301	1.23	0.194	0.90-1.69
P for trend			1.08 [†]	0.143	
Mean density, % (SD)	24.8 (8.1)	24.2 (7.2)			
Interval cancer population					
Wolfe's patterns					
N1	24	170	1 (reference)	—	—
P1	34	151	1.55	0.134	0.87-2.77
P2	110	252	3.11	<0.001	1.90-5.10
DY	38	52	4.80	<0.001	2.58-8.93
P for trend			1.74 [†]	<0.001	
Visual percent density					
0-15%	38	239	1 (reference)	—	—
20-35%	26	120	1.40	0.240	0.80-2.43
40-70%	82	175	2.99	<0.001	1.93-4.55
75-90%	60	91	4.07	<0.001	2.48-6.70

(Continued on the following page)

Table 2. Association between breast density and risk in the overall, screen-detected, and interval cancer populations (Cont'd)

	Cases (n)	Controls (n)	OR*	P	95% CI
<i>P</i> for trend			1.65 [†]		
Mean density, % (SD)	51.8 (27.7)	34.5 (28.3)			
Interactive threshold			1 (reference)	—	—
0.00-9.75%	20	188	2.30	0.006	1.28-4.16
9.76-23.62%	39	169	4.14	<0.001	2.35-7.29
23.63-38.52%	60	148	7.31	<0.001	4.17-12.8
38.63-82.28%	87	120	1.89 [†]	<0.001	
<i>P</i> for trend					
Mean density, % (SD)	35.7 (19.7)	22.3 (17.6)			
SMF			1 (reference)	—	—
11.6-19.4%	26	182	2.22	0.004	1.28-3.82
19.5-22.8%	46	162	2.54	0.001	1.49-4.36
22.9-28.4%	52	157	5.77	<0.001	3.40-9.79
28.5-55.6%	82	124	1.71 [†]	<0.001	
<i>P</i> for trend					
Mean density, % (SD)	27.2 (7.5)	23.6 (6.8)			

*Adjusted for age, mode of diagnosis, and screening center.

[†] ORs for every increase in respective density quartiles or categories (Wolfe's patterns).

the threshold method compared with Wolfe's classification, visual estimate of percent density, and the SMF method (Table 2). The ORs determined by density measures using the threshold were 2.45 (95% CI, 1.86-3.23), for the SMF method was 1.92 (95% CI, 1.47-2.51), for Wolfe's grades was 1.79 (95% CI, 1.27-2.52), and for visual percent density scoring was 1.81 (95% CI, 1.41-2.32), comparing the highest categories of density with the baseline category.

Table 2 summarizes the ORs associated with increasing categories of density in separate screen-detected and interval cancer populations. In the screen-detected cancer group, Wolfe's grading, visual percent density scoring, and SMF measures of percent density were not related to cancer risk, as shown by the large *P* values, although there was some suggestion of a stepwise increase in risk with density (Table 3). Density measure as estimated by the threshold method was significantly related to cancer risk (*P* for trend = 0.007), with ORs of 1.56 (95% CI, 1.13-2.17) comparing highest with baseline density quartiles. In the interval cancer

population, the difference in mean percent density between cancers and controls was much greater for the threshold measures compared with SMF measures (Table 2). Density estimates as measured using all four methods were significantly associated with increased cancer risk and the observed ORs were two to three times higher than those in the screen-detected group (Table 2). Cancer risk for threshold method comparing highest with lowest density categories for the threshold method was 7.31 (95% CI, 4.17-12.8), for SMF method was 5.77 (95% CI, 3.40-9.79), for Wolfe's patterns was 4.80 (95% CI, 2.58-8.93), and for visual density scoring was 4.07 (95% CI, 2.48-6.70).

There is variation in the association of density with cancer based on the method used to estimate density. For example, a woman with ~40% area of density or more in her breast would be at 2.45-fold risk according to the threshold method but only 1.24- to 1.81-fold risk according to the visual grading (Table 2). The distribution of SMF results is also different from threshold readings. For example, a woman whose breast is

Table 3. Density-risk association adjusted for interactive threshold and SMF

	Cases (n)	Controls (n)	OR*	P	95% CI
Interactive threshold [†]			1 (reference)	—	—
0.00-9.41%	112	517	1.39	0.027	1.04-1.86
9.43-22.92%	144	484	1.78	<0.001	1.30-2.45
22.94-37.03%	171	458	2.24	<0.001	1.58-3.18
37.04-84.74%	207	421	1.30 [‡]	<0.001	1.17-1.46
Continuous measure	634	1,880			
SMF [§]			1 (reference)	—	—
11.5-19.1%	137	500	0.94	0.672	0.71-1.25
19.2-22.7%	142	492	0.92	0.608	0.67-1.26
22.8-28.2%	155	461	1.16	0.382	0.83-1.64
28.3-59.3%	200	427	1.05 [‡]	0.407	0.94-1.17
Continuous measure	634	1,880			

*Adjusted for age, mode of diagnosis, and screening center.

[†] Adjusted for SMF results.

[‡] ORs for every increase in respective density quartile.

[§] Adjusted for interactive threshold readings.

predominantly fatty and is classified as having virtually no density measured using the threshold method will have a minimum of 11.5% density based on the calculations of SMF. This is because in the computation of SMF, even the smallest amount of density must be stretched out across the thickness of the breast, so there are rarely very low volumetric percent density values.

To determine if lack of calibration data affected the observed risk, we eliminated the set of 124 films with their matched cases/controls and carried out the same multivariate analyses. A total of 558 cases and 1,804 controls were used in this subanalysis. The ORs obtained were consistently lower for all methods, and the absolute decreases in OR were no more than 12% (data not shown; available on request). The results that were significant in the total data set were also significant in this subset of data.

We tested the effect of different digitizers on SMF results by directly comparing the SMF values for Lumisys-scanned images and those for Array-scanned images. We showed a 100% correlation in the SMF results (18).

Using the likelihood ratio, the model fit for the four methods was compared. The χ^2 statistic for SMF measures, threshold readings, Wolfe's grades, and visual readings was 5.93, 25.42, 2.86, and 3.47, respectively, suggesting that the threshold readings provided the best predictors of breast cancer risk.

After adjusting for interactive threshold readings, SMF density results were no longer associated with breast cancer risk (Table 3). However, adjustment for SMF results did not significantly modify the relationship between threshold density readings and cancer risk (Table 3) compared with before the adjustment (Table 2).

Discussion

Main Findings. To our knowledge, this is the first study to evaluate the potential of the Generate SMF program to predict breast cancer risk. Our data suggest that the implementation of SMF that was commercially available and used in this project does not currently provide as much information on breast cancer risk as the widely used interactive threshold software. Several lines of evidence in this study support this conclusion. First, the stepwise increase in density-risk association was greater using the threshold method of density measurement compared with the other methods, including SMF. In other words, the density estimates from threshold method capture a relatively better overall density-risk association. Furthermore, only this measure of density was related to breast cancer risk in the screen-detected cancer population. The likelihood ratio test also showed that the thresholding technique was the best method in the model fit. Finally, density measures from the threshold method convey more information on breast cancer risk than does the SMF program. Once threshold readings of percent density are taken into account, no more information can be obtained by the SMF density measures.

Mammographic features captured by the threshold method have the strongest association with cancer risk compared with Wolfe's patterns, visual estimation of

percent density, and SMF. The comparison among the three remaining methods was more ambiguous. Although the stepwise increase for density and risk association was overall greater for the SMF method, the differences were too small for us to form definitive conclusions. The likelihood ratio test provided further evidence that SMF was the best predictor of risk of the three methods. However, more work is needed to explore this conjecture.

It is generally accepted that the association between density and risk is stronger for percent density measures compared with Wolfe's patterns or other qualitative classifications (1-5, 19, 20). Most studies arrived at these conclusions based on the relative magnitudes of risk or ORs. There are several limitations in such studies, including inconsistencies in the methods used to categorize percent density, differences in the experience of the assessors, and variations in study populations. To overcome these limitations, some articles carried out direct comparisons between qualitative and quantitative methods and their associations with cancer risk (3, 21-23). Our findings are partly in agreement with the work by Brisson et al. (3), showing percent density as a stronger risk predictor compared with Wolfe's grading. However, their percent density readings were obtained visually using a method similar to the 21-scale system in our study.

The magnitudes of breast cancer risk relative to each categories of density were lower than in most other studies. This may be due to differences in the range of readings between studies (e.g., a wider range of readings results in the upper density category being farther from the baseline group, leading to a larger observed risk). Furthermore, the lower than average ORs can be explained by a lack of confounder data, such as body mass index, family history of breast cancer, and reproductive variables, all of which are strongly related to breast cancer risk. Stratifying the data by mode of detection revealed that Wolfe's patterns, visual percent density, and SMF percent density were not associated at all with cancer risk in the screen-detected population. This finding is not in accordance with results from several other studies (2, 3, 5, 7, 21). Relative risks with respect to density were severalfold higher in the interval cancer population than the screen-detected group. This could be explained by the masking hypothesis that states that tumors can be masked in the midst of the extensive breast density and are thus easily missed during routine screening but are presented symptomatically after a negative screen (24). Consequently, women with dense breasts are at higher risk of developing interval cancers compared with screen-detected cancers.

Three of the four methods in the study have a subjective component, which is the main limitation of their use. Although SMF has perfect repeatability, the accuracy of its density measures may rely on the amount of X-ray variable information available. However, the calibration variable compensation mechanism of SMF underwent evaluation and showed robustness in estimating breast thickness even in the absence of any calibration data (16). Finally, we must emphasize that there is a need to be able to apply SMF to digital mammograms, for which the present software is not adapted.

Several groups around the world have been working on developing volume-based density measurement methods. Most methods use a step-wedge approach to calibrate each image, which must be carried out at the time of mammography and therefore cannot be applied retrospectively to films (25). Some teams are also working on a new version of the threshold method applied to digital mammography. Advances in this subject will be of great value as the area of mammography is quickly moving toward digital forms. There are advantages to using digital images over film mammography, including lack of labor for film digitizing, lack of differences in the quality of digitized images, and availability of calibration data for digital images.

Strengths and Limitations. Despite subjectivity in the three methods evaluated in this study, the films were all assessed by the same reader who has shown good repeatability. Furthermore, this reader was trained by the originator of the threshold technique (Professor Norman Boyd, Toronto, Ontario, Canada) to ensure good interclass agreement on using this method. To determine the ability of SMF to predict breast cancer risk, its measures of density were compared with two of the most widely used density assessment methods, Wolfe's grading and interactive threshold. The large size of the study population and the high ratio of controls to cancers provided further statistical robustness to this case-control study. Finally, ~95% of the population data sets have full film acquisition data, which could be important for the accuracy of SMF.

The main limitation in this study is the lack of confounding data, as this was a retrospective study involving medical records only. However, this lack of adjustment for confounders should theoretically affect all three manual reading methods equally. For example, we have evidence (7, 21, 26, 27) that body mass index is a confounder for the Wolfe's grading and the threshold method, but we do not know whether this applies also to SMF. Furthermore, the quartile method of categorizing the continuous measures of density was not consistently used in the previous studies, so the magnitudes of breast cancer risk could not be fairly compared with those in other studies.

Future Directions. Breast density is currently undergoing investigation as a biomarker for breast cancer risk. Although the current implementation of GeneRateSMF is not the best predictor of breast cancer risk, several teams are working on improving the SMF software and possibly applying the new algorithm on digital mammography. Future work on producing an objective volumetric approach of measuring breast density will be a step closer to the use of density in clinical practice and its involvements in studies on breast cancer etiology.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be

hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank the staff at the Cambridge Breast Screening Unit, particularly Dr. Peter Britton, Dr. Lynda Bobrow, Judith Fatibene, and Barbara Knighton; the staff at the Norwich Breast Screening Unit, with special mentions to Dr. Erika Denton, Dr. Anne Girling, Glynis Wivell, and Judy Priddle; the technical team at Siemens Molecular Imaging, including Jerome Declerck and Xiao Pan Bo; Professor Norman Boyd for his training on using the Cumulus program; and Professor Martin Yaffe and his team for technical support.

References

1. Wolfe JN, Saftlas AF, Salane M. Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study. *Am J Roentgenol* 1987;148:1087–92.
2. Brisson J, Verreault R, Morrison AS, Tennina S, Meyer F. Diet, mammographic features of breast tissue, and breast cancer risk. *Am J Epidemiol* 1989;130:14–24.
3. 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;12:728–32.
4. Saftlas AF, Hoover RN, Brinton LA, et al. Mammographic densities and risk of breast cancer. *Cancer* 1991;67:2833–8.
5. Byrne C, Schairer C, Wolfe J, et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. *J Natl Cancer Inst* 1995;87:1622–9.
6. Boyd NF, Byng JW, Jong RA, et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. *J Natl Cancer Inst* 1995; 87:670–5.
7. Ursin G, Ma H, Wu AH, et al. Mammographic density and breast cancer in three ethnic groups. *Cancer Epidemiol Biomarkers Prev* 2003;12:332–8.
8. Ciatto S, Bravetti P, Cecchini S, Cirillo A. Mammographic parenchymal patterns and breast cancer risk. A case control study of a population-based screening experience. *Radiol Med (Torino)* 1990;79: 346–8.
9. Wolfe JN. Breast patterns as an index of risk for developing breast cancer. *Am J Roentgenol* 1976;126:1130–7.
10. Cuzick J, Warwick J, Pinney E, Warren RM, Duffy SW. Tamoxifen and breast density in women at increased risk of breast cancer. *J Natl Cancer Inst* 2004;96:621–8.
11. Byng JW, Yaffe MJ, Jong RA, et al. Analysis of mammographic density and breast cancer risk from digitized mammograms. *Radiographics* 1998;18:1587–98.
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;79:378–82.
13. McCormack VA, Highnam R, Perry N, dos Santos Silvia I. Comparison of a new and existing method of mammographic density measurement: intramethod reliability and associations with known risk factors. *Cancer Epidemiol Biomarkers Prev* 2007;16: 1148–54.
14. Byng JW, Yaffe MJ, Jong RA, et al. Symmetry of projection in the quantitative analysis of mammographic images. *Eur J Cancer Prev* 1996;5:319–27.
15. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. *Phys Med Biol* 1994;39: 1629–38.
16. 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: 2695–713.
17. Highnam R, Brady J. *Mammographic image analysis*. Dordrecht (the Netherlands): Kluwer Academic; 1999.
18. Ding J. Case control study on the effectiveness of using standard mammogram form to predict breast cancer risk [dissertation]. Cambridge (UK): University of Cambridge; 2007.
19. McCormack VA, dos Santos Silvia I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;15:1159–69.
20. Warner E, Lockwood G, Trichtler D, Boyd NF. The risk of breast cancer associated with mammographic parenchymal patterns: a meta-analysis of the published literature to examine the effect of method of classification. *Cancer Detect Prev* 1992;16:67–72.

21. Torres-Mejia G, De Stavola B, Allen DS, et al. Mammographic features and subsequent risk of breast cancer: a comparison of qualitative and quantitative evaluations in the Guernsey prospective studies. *Cancer Epidemiol Biomarkers Prev* 2005;14:1052–9.
22. Thomas DB, Carter RA, Bush WH, Jr., et al. Risk of subsequent breast cancer in relation to characteristics of screening mammograms from women less than 50 years of age. *Cancer Epidemiol Biomarkers Prev* 2002;11:565–71.
23. Gram IT, Bremnes Y, Ursin G, Maskarinec G, Bjurstam N, Lund E. Percentage density, Wolfe's and Tabar's mammographic patterns: agreement and association with risk factors for breast cancer. *Breast Cancer Res* 2005;7:R854–61.
24. van Gils CH, Otten JD, Verbeek AL, Hendriks JH. Mammographic breast density and risk of breast cancer: masking bias or causality? *Eur J Epidemiol* 1998;14:315–20.
25. Pawluczyk O, Augustine BJ, Yaffe MJ, et al. A volumetric method for estimation of breast density on digitized screen-film mammograms. *Med Phys* 2003;30:352–64.
26. Gill JK, Maskarinec G, Pagano I, Kolonel LN. The association of mammographic density with ductal carcinoma *in situ* of the breast: the Multiethnic Cohort. *Breast Cancer Res* 2006;8:R30.
27. Maskarinec G, Pagano I, Lurie G, Wilkens LR, Kolonel LN. Mammographic density and breast cancer risk: the multiethnic cohort study. *Am J Epidemiol* 2005;162:743–52.

Cancer Epidemiology, Biomarkers & Prevention

Evaluating the Effectiveness of Using Standard Mammogram Form to Predict Breast Cancer Risk: Case-Control Study

Jane Ding, Ruth Warren, Iqbal Warsi, et al.

Cancer Epidemiol Biomarkers Prev 2008;17:1074-1081.

Updated version Access the most recent version of this article at:
<http://cebp.aacrjournals.org/content/17/5/1074>

Cited articles This article cites 25 articles, 7 of which you can access for free at:
<http://cebp.aacrjournals.org/content/17/5/1074.full#ref-list-1>

Citing articles This article has been cited by 8 HighWire-hosted articles. Access the articles at:
<http://cebp.aacrjournals.org/content/17/5/1074.full#related-urls>

E-mail alerts [Sign up to receive free email-alerts](#) related to this article or journal.

Reprints and Subscriptions To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions To request permission to re-use all or part of this article, use this link
<http://cebp.aacrjournals.org/content/17/5/1074>.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.