The Detection of Changes in Mammographic Densities

Giske Ursin,[1] Melvin A. Astrahan, Martine Salane,
Yuri R. Parisky, John G. Pearce, John R. Daniels,
Malcolm C. Pike, and Darcy V. Spicer

Departments of Preventive Medicine [G. U., M. C. P.], Radiation Oncology
[M. A. A.], Radiology [Y. R. P., J. G. P.], and Medicine [D. V. S., J. R. D.],
University of Southern California/Norris Comprehensive Cancer Center, Los
Angeles, California 90033, and J. Wolfe Radiological Services P.C., Detroit,
Michigan 48201 [M. S.]

Abstract

We previously reported reductions in mammographic
densities in women participating in a trial of a
gonadotropin-releasing hormone agonist (GnRHA)-based
regimen for breast cancer prevention. In our previous
report, we compared (by simultaneous evaluation) three
basic elements of mammographic densities. The purpose
of the present study was to evaluate whether a standard
(expert) method of measuring mammographic densities
would detect such changes in densities and whether a
novel nonexpert computer-based threshold method could
do so.

Mammograms were obtained from 19 women at
baseline and 12 months after randomization to the
GnRHA-based regimen. The extent of mammographic
densities was determined by: (a) a standard expert
outlining method developed by Wolfe and his colleagues
(Am. J. Roentgenol., 148: 1087–1092, 1987); and (b) a
new computer-based threshold method of determining
densities.

The results from both the expert outlining method
and the computer-based threshold method were highly
consistent with the results of our original (simultaneous
evaluation) method. All three methods yielded statistically
significant reductions in densities from baseline to the
12-month follow-up mammogram in women on the
contraceptive regimen. The difference between the
treated and the control group was statistically significant
with the expert outlining method and was of borderline
statistical significance with the computer-based threshold
method. The computer-based results correlated highly
($r > 0.85$) with the results from the expert outlining
method.

Both the standard expert outlining method and the
computer-based threshold method detected the reductions
we had previously noted in mammographic densities
induced by the GnRHA-based regimen.

Introduction

Wolfe (1, 2) originally described four parenchymal patterns
(N1, P1, P2, and DY) of increasing mammographic densities
and found that the risk of incident breast cancer increased with
increasing density pattern and was much greater in women with
the DY pattern than in women with the N1 pattern (2). Other
investigators reported similar success with systems in which the
reader simply estimated the proportion of the breast containing
glandular or ductal densities from the mammogram (3–5).

Wolfe *et al.* (6) subsequently described a method of ac-
tually measuring the extent of mammographic densities. In this
method, experienced readers (experts) outline the area with
densities directly on the mammogram. The area containing
densities and the total area of the breast are subsequently
measured with a planimeter, and the percentage of the breast
containing densities is calculated. The percentage of the breast
containing densities determined by this method has been found
to be highly associated with risk of breast cancer (6–8).

Byng *et al.* (9) have developed a computer-based threshold
method for measuring densities. Mammograms are digitized
using a high-resolution laser scanner, and the scanned image is
displayed on a computer screen. The reader selects a threshold
value (gray scale on the computer screen) that best distin-
guishes the breast from the dark background. An edge detection
computer program (algorithm) is used to outline the breast, and
the computer measures the total area of the breast (in terms of
computer screen pixels). Subsequently, the reader identifies a
second threshold value, the gray value that best identifies the
edges of the mammographically dense areas within the breast
parenchyma. The number of pixels in the areas containing such
densities is then measured by the computer, and the fraction of
pixels in the breast parenchyma that are located in the dense
areas (percentage of mammographic densities) is calculated.
This computer-based threshold method has been found to be
equally reproducible for experienced and inexperienced readers
(9) and to be strongly predictive of breast cancer risk (10).

Our goal was to develop a method of determining mam-
mographic densities that could be used in epidemiological
studies by personnel with minimal or no radiology training and
that could be implemented on relatively reasonable equipment.
We report here our development of a computer-based system
that we adapted from the system of Byng *et al.* (9). This has
been implemented on a standard personal Apple Power Macin-
tosh computer with a relatively low-resolution scanner.

For our pilot trial of a GnRHA[2]-based breast cancer che-
moprevention regimen, we developed a comparative method in
which the mammographic densities in mammograms obtained
before and after use of the regimen could be compared by
simultaneous evaluation (11).

In the current project, both the expert outlining method (6)
and our computer-based threshold method were applied to the

mammograms from the GnRHA trial (11). The primary purpose of this was to evaluate whether a more standard method of measuring mammographic densities (the expert outlining method) would also detect the changes noted with our nonstandard simultaneous evaluation method. A second purpose was to evaluate whether the standard expert outlining method could be replaced by a nonexpert computer-based threshold method.

Patients and Methods

Mammograms. Women in this study were participants in a clinical trial of a GnRHA-based regimen designed to reduce breast cancer risk (11). Twenty-one women ages 25–40 years were randomly assigned in a 2:1 ratio to the contraceptive regimen (14 women) or to a control group (7 women). One woman was later removed from the contraceptive group because of poor compliance with the regimen. A second woman in the contraceptive group had breast implants, and her mammograms were found unsuitable for inclusion in this study for technical reasons. Women in the contraceptive group received: (a) 7.5 mg of leuprolide acetate depot (Lupron Depot) by i.m. injection every 28 days; (b) 0.625 mg of oral conjugated estrogen (Premarin) 6 of 7 days every week; and (c) 10 mg of oral medroxyprogesterone acetate (Provera) for the last 13 days of every fourth 28-day cycle. The regimen was designed to minimize exposure of the breast epithelium to estrogen and progestin while preserving the beneficial effects of estrogen on cardiovascular disease risk and still preventing endometrial hyperplasia.

Mammograms were obtained at baseline and after 12 months on study. The craniocaudal mammograms were used to study mammographic changes from baseline to 12 months in the treated and control groups. Four mammograms per woman were used (right and left breasts at baseline and 12 months). All readings of mammograms (by all three methods described below) were done masked as to baseline or 12-month status and as to being from treated or control women.

Simultaneous Evaluation Method. This is the method we originally used (11). Each baseline film was matched with the corresponding 12 month film. The films in each pair were designated A and B at random. Each pair was then presented to two experienced radiologists (Y. R. P. and J. G. P.) familiar with reading mammograms. Each pair of films was independently compared by each reader in terms of: (a) clarity of fibrous septae (trabeculae), in which greater clarity is associated with less density; (b) amount of nodular collections of glandular tissue; and (c) amount of confluent areas, in which breast tissue is arranged as diffuse sheets. The results were recorded on a 5-point scale (−2, −1, 0, 1, and 2) as a difference between the 12 month and baseline mammograms, in which a negative value indicated a reduction in density from baseline to 12 months. The individual subject's average change scores (both views) for each reader were calculated for all three questions combined. For the comparison between this method and the other two methods, the results from only one of the radiologists were used. (The results were essentially unaffected by the particular choice of radiologist.)

Expert Outlining Method. One of us (M. S., who is very experienced with the method of outlining densities; Refs. 6–8) outlined the area of the breast that contained mammographic densities on each of the 76 mammograms. Each mammogram (with the outline) was subsequently digitized, and the outlined area of the breast containing densities was measured (effectively in terms of computer screen pixels). Finally, the total area of

the breast was outlined on the digitized image by one of us and measured (see below for details). After estimating the percentage of the breast with densities separately for the left and the right breast, the average percentage densities for the two breasts was calculated.

For the expert outlining method, two measures of effect were calculated: (a) the difference between the 12-month and baseline measures of the percentage of the breast containing densities (percentage densities), expressed as a percentage of the baseline measure; and (b) the difference between the 12-month and baseline density areas (absolute densities), again expressed as a percentage of the baseline measure.

Computer-based Threshold Method. All mammograms were digitized using an Omnimedia XRS 6cx scanner (Lumisys, Sunnyvale, CA). This scanner creates an 8-bit (256 shades) gray scale image that is linear in the absorbance range of 0–2.8. The mammograms were scanned at a resolution of 150 pixels/inch (59 dots/cm). A pixel value of 0 represents the darkest (black) shade in the image; a value of 255 represents the lightest (white) value.

The digitized images were analyzed using a Macintosh Operating System (MacOS)-compatible personal computer system consisting of an Apple Power Macintosh 7100/66 and an Apple Multiple Scan 17 video monitor. The video display was set for 24-bit color mode at a resolution of 832 × 624 pixels. In 24-bit color mode, each screen pixel is represented and stored as three 8-bit fields, one each for RGB. Gray is achieved by setting the three RGB fields to the same value. The 8-bit gray scale images were converted to equivalent 24-bit images for display purposes by setting each of the RGB fields of a 24-bit pixel equal to the value of the corresponding 8-bit pixel.

The software used for measuring densities was developed by one of us (M. A. A.). The program provides numerous tools for image processing and interactively defining regions of interest. One tool instantaneously applies a yellow tint to any interactively selected subrange of a gray scale image. The yellow tint is applied nondestructively by setting the red field of each RGB pixel to 0 while leaving intact the green and blue fields. The original gray is restored by simply resetting the red field equal to either the green or blue field (because all three were originally equal).

On the digitized mammographic image displayed on the screen, the reader first defines a ROI for analysis; this region includes all of the breast shown on the mammogram but specifically excludes the pectoralis muscle (same area used for Wolfe's method). The reader then uses the tinting tool to apply a yellow tint to gray levels above some threshold X (*i.e.*, pixels corresponding to 8-bit gray levels $\geq X$ and ≤ 255). The reader searches for the best threshold in which all pixels $\geq X$ are considered to represent mammographic densities. The software counts both the total number of pixels and the number of tinted pixels within the defined ROI. The fraction (percentage) of the breast with densities is taken as the ratio of the tinted area:total area of ROI. After estimating the percentage of the breast with densities separately for the left and the right breast, the average percentage densities for the two breasts is calculated.

This method was implemented by one of us (G. U.), a medically trained epidemiologist who first underwent a training session with M. S. 6 months before conducting the actual readings. In this session, M. S. described the expert outlining method of identifying densities and illustrated the method using 14 of 76 mammograms. G. U. outlined an additional three mammograms under the critique of M. S. She carried out all of the computer-based readings for the study 6 months after com-

Table 1   Changes in mammographic densities between baseline and 12 months in women treated with a GnRHA-based contraceptive regimen and controls assessed by three different methods

Group	Person no.	Simultaneous evaluation method (change score)[a]	Expert outlining method (% change in % densities)	Computer-based threshold method (% change in % densities)[b]
Treated	1	−1.67	−38.0	−38.1
	2	−1.33	−27.9	−37.1
	3	−1.33	−38.9	−8.3
	4	−1.00	−70.8	−79.1
	5	−1.00	−39.1	−31.4
	6	−0.92	4.7	−24.1
	7	−0.67	−34.7	−39.5
	8	−0.58	−30.2	−44.4
	9	−0.58	−22.5	17.6
	10	−0.17	−2.3	8.2
	11	0.25	−13.7	−0.9
	12	0.67	−0.1	18.2
Mean (±SE)		−0.69 (± 0.20)	−26.12 (± 6.10)	−21.6 (± 8.43)
P for change[c]		0.010	0.002	0.027
Control	1	−0.92	−44.2	−21.2
	2	−0.50	−21.8	24.7
	3	−0.50	28.1	−1.2
	4	0.08	−17.2	−19.4
	5	0.17	3.8	15.7
	6	0.50	33.0	−18.0
	7	1.00	38.2	30.9
Mean (±SE)		−0.05 (± 0.25)	2.83 (± 31.67)	−1.64 (± 8.36)
P for change[c]		0.94	0.81	0.94
P for difference between groups[d]		0.042	0.069	0.083

[a] Only data from one of the radiologists are displayed.
[b] Data from the first of three readings.
[c] *P*s for within group change were calculated using Wilcoxon's signed rank test.
[d] *P*s for differences between the treated and control group were calculated using Wilcoxon's rank-sum test.

pleting the training session. The mammograms were read three times. For the comparison between the methods described here, the results from the first of these three readings were used.

Statistical Methods. In the tables, we show the difference in density score at 12 months minus the density score at baseline, using both differences in absolute areas and differences in the percentage of breast with densities.

The statistical significance of the difference between the 12 month and baseline mammograms within each group (GnRHA regimen and control) was evaluated using Wilcoxon's nonparametric signed ranks test (12). The statistical significance of the differences in the mammographic changes between the contraceptive group and the control group was evaluated using Wilcoxon's nonparametric rank-sum test (12).

Pearson's correlation coefficients were calculated between the expert outlining method and the computer-based threshold method and between the multiple readings performed by the same person using the computer-based threshold method.

Statistical significance levels (*P*) quoted are two-sided and are written as 2*P*.

Results

Simultaneous Evaluation Method. With the simultaneous evaluation method, there was a reduction in densities in 10 of 12 women in the treatment group after 12 months, compared to a reduction in 3 of 7 control women (Table 1). The largest reductions were almost exclusively concentrated in the GnRHA-based regimen group. With this restricted data set and using only one reader, the difference between baseline and 12 month mammograms was statistically significant in the treatment group (2*P* = 0.010). The change in the treatment group

was also statistically significant when compared with the change in the control group (2*P* = 0.042).

Expert Outlining Method. With the expert outlining method, there was a reduction in percentage densities in 11 of 12 women in the treatment group after 12 months, compared to a reduction in 3 of 7 control women, and the largest reductions were again almost exclusively concentrated in the GnRHA-based regimen group. The reduction in densities in the treatment group was highly statistically significant (2*P* = 0.002), but the difference between the treatment group and the control group did not quite reach conventional statistical significance (2*P* = 0.069).

The results when evaluating differences in absolute densities estimated with the expert outlining method were very similar to the results when evaluating differences in percentage densities (data not shown). There was a statistically significant reduction in densities between baseline and 12 month mammograms in the treatment group (2*P* = 0.012), and the change in the treatment group was statistically significant when compared with the change in the control group (2*P* = 0.025).

Computer-based Threshold Method. The computer-based threshold method showed a reduction in percentage densities in 9 of 12 women in the treatment group after 12 months, compared to a reduction in 4 of 7 control women, and the largest reductions were again almost exclusively concentrated in the GnRHA-based regimen group. The reduction in percentage densities in the treatment group was statistically significant (2*P* = 0.027); but the difference between the treatment group and the control group did not quite reach conventional statistical significance (2*P* = 0.083).

As with the expert outlining method, the results were similar for absolute densities.

Table 2 Correlations between the expert outlining method and the computer-based threshold method on 76 mammograms

	Computer-based threshold method reading number:		
	1	2	3
Expert outlining method	0.86	0.89	0.91
Computer-based threshold method reading number:			
1		0.96	0.92
2			0.95

Correlations between the Three Different Methods.

The correlations between the simultaneous evaluation method and the expert outlining method were 0.68 (percentage densities) and 0.76 (absolute densities). The correlations between the results from the simultaneous evaluation method and the computer-based threshold method were 0.62 (percentage densities) and 0.66 (absolute densities).

The correlations between the difference results from the expert outlining method and the computer-based threshold method were 0.64 for percentage densities and 0.70 for absolute densities. The correlation coefficients between these two methods for measuring densities (not differences) for all of the 76 mammograms are shown in Table 2. There was a high correlation between the two methods, and the three readings made with the computer-based threshold method were also highly correlated.

Discussion

Both the expert outlining method of Wolfe *et al.* (6) and our computer-based threshold method showed high consistency with the previous method we had used to distinguish baseline mammographic densities from mammographic densities after 12 months on the GnRHA-based regimen (11). The three methods gave very similar results with no clear difference in statistical power between our original simultaneous evaluation method and Wolfe's (expert) outlining method. There may be a slight loss of statistical power with the computer-based (non-expert) system.

We have not presented data on intra- and interobserver reliability for the simultaneous evaluation method. Extensive use of this method would warrant collection of such data.

Both the expert outlining method and our computer-based threshold method clearly detected changes in mammographic densities.

The results from the computer-based threshold method correlated highly with those using the expert outlining method. This is in accordance with what Byng *et al.* (9) found with their computer-based threshold method.

For epidemiological purposes, an advantage of the computer-based threshold method is that once the images have been digitized, they can be read at any time, one or multiple times, in whichever order desired. Furthermore, the computer-based threshold method requires less training than both the expert outlining method and the simultaneous evaluation method.

The epidemiologist using the computer-based threshold method was trained on a subset of the 76 mammograms that were later used in the study. The number of mammograms and the 6-month time lapse since training made it most unlikely that the previous training film results influenced the subsequent real reading results of these mammograms. Furthermore, the correlation between the expert outlining method and the computer-

based threshold method was high and could not be explained on the basis of a small subset of the mammograms. It is thus most unlikely that the results shown here are materially affected by this.

Although we have not demonstrated that our computer-based threshold method can predict breast cancer risk in a population-based study, the high correlations with the expert outlining method suggest that it will be similarly effective in predicting breast cancer risk.

The expert outlining method has been strongly associated with breast cancer risk in previous studies. Saftlas *et al.* (7) used the expert outlining method on craniocaudal mammograms to measure the percentage of the breast containing densities. Categorizing percentage densities into approximate quintiles, they found that successive quintiles had relative risks of 1.7, 2.5, 3.8, and 4.3 ($2P < 0.0001$). Similarly, applying the expert outlining method to data from the Breast Cancer Detection Demonstration Project, Byrne *et al.* (8) found that compared with women who had no mammographic densities, women with 75% or more densities had a more than 4-fold elevated risk of breast cancer (95% confidence interval, 3.1–6.1).

Boyd *et al.* (10) used the computer-based threshold method of Byng *et al.* (9) and found that women with 75% or more densities had a 4-fold elevated risk of breast cancer compared to women with no densities ($2P = 0.0001$).

The results presented here for the computer-based system were from a single reader. We had two naïve readers conduct mammographic densities assessments after only minimal training on 10 mammographic images. Despite this minimal exposure to mammographic density readings, these two observers obtained correlation coefficients with the expert outlining method on the 76 mammograms of approximately 0.7. One of the problems that these naïve readers had was to compensate for differences in quality and exposure of the mammogram. This was easier for more experienced readers. However, Byng *et al.* (9) have reported that a naïve observer can perform as well as a highly trained one. We are currently in the process of training other readers on an extensive set of films.

Technical Issues. In this study, all mammograms were scanned on a scanner with 8-bit pixel depth ($2^8 = 256$ shades of gray) and with a resolution of 150 dpi. Standard personal computer monitors can only display 256 levels of gray; if images are scanned at a higher pixel depth, the computer program must convert these down to 256 to display the mammographic image. With the type of monitor we used, a 17-inch monitor with pixel size of 0.28 mm (which corresponds to a resolution of approximately 90 dpi), scanning the images at a resolution of 150 dpi seems adequate. We did, however, also scan a subset of the images with a higher-resolution scanner, a Cobrascan CX-312T scanner (Radiographic Digital Imaging, Inc., Compton, CA). This scanner provides a 12-bit pixel depth (4096 shades of gray) at a resolution of 300 dpi. This did not improve the computer images when displayed on a conventional Apple monitor.

Significance of Mammographic Density Reductions. There is substantial epidemiological and experimental evidence that ovarian hormones (in particular, estrogens, and possibly progesterone) increase the risk of breast cancer (13–15). The GnRHA-based regimen attempts to reduce the levels of estrogen and progestogen to a minimum while still preserving the essential beneficial effects of estrogen. Blocking of ovarian function is achieved through the use of the GnRHA; sufficient estrogen is given to prevent hypoestrogenic symptoms (such as hot flashes), and intermittent progestogen is given to prevent

any estrogen-induced endometrial hyperplasia. The fact that mammographic densities are reduced after a year on this regimen suggests that such a regimen may protect against breast cancer (11). In that regard, it is interesting to compare the density reduction obtained with the GnRHA regimen with that obtained in a 2-year randomized trial of a low-fat, high-carbohydrate diet; the GnRHA regimen resulted in a more than three times larger reduction (21.6%) than the diet (6.1%; Ref. 16).

Conclusion. We have demonstrated that both the standard expert outlining method and our new computer-based threshold method detected the mammographic density reductions induced by the GnRHA-based regimen. The computer-based threshold method of determining mammographic densities was found to be highly reproducible and correlated well with the expert outlining method. The possibility of using this computer-based threshold method for evaluating the effects of other preventive regimens should be explored further.

Acknowledgments

References

1. Wolfe, J. N. Breast patterns as an index of risk for developing breast cancer. Am. J. Roentgenol., *126:* 1130–1139, 1976.

2. Wolfe, J. N. Breast parenchymal patterns and their changes with age. Radiology, *121:* 545–552, 1976.

3. Threatt, B., Norbeck, J. M., Ullman, N. S., Kummer, R., and Roselle, P. Association between mammographic parenchymal pattern classification and incidence of breast cancer. Cancer, *45:* 2550–2556, 1980.

4. Boyd, N. F., O'Sullivan, B., Campbell, J. E., Fishell, E., Simor, I., Cooke, G., and Germanson, T. Mammographic signs as risk factors for breast cancer. Br. J. Cancer, *45:* 185–193, 1980.

5. Brisson, J., Merletti, F., Sadowsky, N. L., Twaddle, J. A., Morrison, A. S., and Cole, P. Mammographic features of the breast and breast cancer risk. Am. J. Epidemiol., *115:* 428–437, 1982.

6. Wolfe, J. N., Saftlas, A. F., and Salane, M. Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study. Am. J. Roentgenol., *148:* 1087–1092, 1987.

7. Saftlas, A. F., Hoover, R. N., Brinton, L. A., Szklo, M., Olson, D. R., Salane, M., and Wolfe, J. N. Mammographic densities and risk of breast cancer. Cancer, *67:* 2833–2838, 1991.

8. Byrne, C., Schiarer, C., Wolfe, J., Parekh, N., Salane, M., Brinton, L. A., Hoover, R., and Haile, R. Mammographic features and breast cancer risk: effects with time, age and menopause status. J. Natl. Cancer Inst., *87:* 1622–1629, 1995.

9. Byng, J. W., Boyd, N. F., Fishell, E., Jong, R. A., and Yaffe, M. J. Quantitative analysis of mammographic densities. Phys. Med. Biol., *39:* 1629–1638, 1994.

10. Boyd, N. F., Byng, J., Jong, R., Fishell, E., Little, L., Miller, A. B., Lockwood, G., Tritchler, D., and Yaffe, M. Quantitative classification of mammographic densities and breast cancer risks: results from the Canadian National Breast Screening Study. J. Natl. Cancer Inst., *87:* 670–675, 1995.

11. Spicer, D. V., Ursin, G., Parisky, Y. R., Pearce, J. G., Shoupe, D., Pike, A., and Pike, M. C. Changes in mammographic densities induced by a hormonal contraceptive designed to reduce breast cancer risk. J. Natl. Cancer Inst., *86:* 431–436, 1994.

12. Rosner, B. Fundamentals of Biostatistics, 3rd ed. Boston: PWS-Kent Publishing Co., 1980.

13. Pike, M. C., Spicer, D. V., Dahmoush, L., and Press, M. F. Estrogens, progestogens, normal breast cell proliferation, and breast cancer risk. Epidemiol. Rev., *15:* 17–35, 1993.

14. Bernstein, L., and Ross, R. K. Endogenous hormones and breast cancer risk. Epidemiol. Rev., *15:* 48–65, 1993.

15. Kelsey, J. L., Gammon, M. D., and John, E. M. Reproductive factors and breast cancer. Epidemiol. Rev., *15:* 36–47, 1993.

16. Boyd, N. F., Greenberg, C., Lockwood, G., Little, L., Martin, L., Byng, J., Yaffe, M., and Tritchler, D. Effects at two years of a low-fat high-carbohydrate diet on radiologic features of the breast: results from a randomized trial. J. Natl. Cancer Inst., *89:* 488–496, 1997.

American Association
for Cancer Research

The detection of changes in mammographic densities.

G Ursin, M A Astrahan, M Salane, et al.

Updated version	Access the most recent version of this article at: http://cebp.aacrjournals.org/content/7/1/43

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/7/1/43. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.