Determinants of Percentage and Area Measures of Mammographic Density

Jennifer Stone*, Ruth M. L. Warren, Elizabeth Pinney, Jane Warwick, and Jack Cuzick*  

* Correspondence to Dr. Jennifer Stone, Centre for Molecular, Environmental, Genetic, and Analytic Epidemiology, The University of Melbourne, Level 1, 723 Swanston Street, Carlton, Victoria 3053, Australia (e-mail: stonej@unimelb.edu.au).

Initially submitted March 2, 2009; accepted for publication September 11, 2009.

Mammographic density is one of the strongest predictors of breast cancer risk. Typically expressed as a percentage of the breast area occupied by radiologically dense tissue on a mammogram, its full value may not be realized because of its negative association with body mass index. A simpler measure of mammographic density, independent of other breast cancer risk factors and equally predictive of risk, would be preferable for risk prediction models. Percentage and area measures of mammographic density were determined for 815 women at high risk for breast cancer from the baseline assessments in the International Breast Cancer Intervention Study I, a trial of tamoxifen for breast cancer prevention conducted between 1992 and 2001. Multivariate linear regression was used to assess associations between risk factors and the mammographic measures. Percent dense area was negatively associated with age, body mass index, menopausal status, predicted risk, and smoking status ($R^2 = 24\%$). Dense area was negatively associated with only age and body mass index ($R^2 = 7\%$), and the latter association was much weaker than for percent dense area. Nondense area was positively associated with age, body mass index, and predicted risk ($R^2 = 36\%$). Dense area was not associated with the multitude of risk factors that percent dense area was, making it a simpler biomarker for risk prediction modeling. Both dense area and percent dense area should be presented whenever possible for comparisons in research.

Abbreviation: SE, standard error.

Women with extensive mammographic density are more likely to develop breast cancer than those of the same age with little or no density (1). Mammographic density is typically expressed as a percentage of the total area of the breast occupied by opaque/white tissue on a mammogram. Denoted here as “percent dense area,” its determinants have been shown to be similar to those associated with breast cancer risk (2) with one notable exception. Increased body mass index has been found to increase a women’s risk of breast cancer (in postmenopausal women) but is negatively associated with percent dense area (3). The latter is likely because a woman’s body mass index is positively related to total breast area, the denominator of percent dense area.

Anthropometric measures such as weight, waist circumference, or body mass index can account for almost a third of the variation in percent dense area (2), and Boyd et al. (3) have reported that failure to adjust for this can lead to an underestimation of the risk estimates for percent dense area. In particular, they found that the strength of association between percent dense area and breast cancer risk increased after adjustment for body mass index. This suggests that the most useful risk factor for breast cancer is not simply percent dense area but, rather, percent dense area adjusted for body mass index (and age).

A simpler measure of density, independent of known breast cancer risk factors and equally predictive of risk, would be highly preferable. It is possible that the absolute area of dense tissue in the breast, instead of the percentage, is less dependent on common risk factors such as body mass index. Moreover, if we are to identify the environmental and genetic determinants of mammographic density in hopes of improving our understanding of the etiology of breast cancer, then it is essential that we investigate the potential separate pathways of disease. The purpose of this paper was to
compare the associations between breast cancer risk factors and the components of percent dense area, absolute dense area and absolute nondense area, separately.

**MATERIALS AND METHODS**

**Subjects and methods**

Subjects and methods for the International Breast Cancer Intervention Study I trial have been described previously and are only briefly mentioned here (4). Baseline mammograms were obtained from 815 women aged 35–70 years at high (at least twice the population) risk of developing breast cancer before being randomly assigned to take either tamoxifen or placebo for 5 years. All of the mammograms were digitized by using a Vidar Advantage Pro digitizer (VIDAR Systems Corporation, Herndon, Virginia) and were also measured (J.S.) by use of Cumulus, a computer-assisted interactive thresholding technique (Imaging Research Program, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Canada) that involves an operator estimating the edge of the breast and setting a threshold level to identify dense tissue. The number of pixels in the digitized image that lie within the defined areas is then automatically calculated and recorded. The result is an absolute measure of the total area of the breast and an absolute measure of dense area, respectively, which when subtracted gives the area of nondense tissue and, when expressed as a ratio, gives a percentage of dense tissue within the breast image.

The films were randomized by subject into reading sets of approximately 100 and measured over the course of 3 weeks. The mediolateral oblique views of the right breasts were used, and the reader was blinded to all other identifying information. A 10% random sample of mammograms was repeated in each set and between every fifth set to test the reliability of the measurement.

**Statistical analysis**

The intraclass correlation coefficient was used to assess the reader’s intra- (within a reading set) and inter- (between reading sets) reliability for each of the computer-assisted measures. Univariate and multivariate linear least-squares regression was used to examine associations between the mammographic measures and known breast cancer risk factors measured at baseline. The predicted risk of each breast cancer variable was generated via the Tyrer-Cuzick model (5) and has been shown to be the most consistently accurate model for prediction of breast cancer in high-risk women (6).

Covariate-adjusted residuals of all of the regression models were inspected for normality, and the baseline measures of dense area and nondense area required square-root transformation. In the text, the regression coefficients of dense area and nondense area were divided by their means (transformed), providing a percentage increase/decrease per risk factor unit on a linear scale.

Fractional polynomial smooth plots were generated in STATA, version 10.1, software (StataCorp LP, College Station, Texas) that calculates the prediction for the mammographic density measure from estimation of a fractional polynomial of body mass index and plots the resulting curve along with the confidence interval of the mean.

**RESULTS**

Baseline mammograms from 815 women from the International Breast Cancer Intervention Study I were obtained; however, 11 mammograms were excluded because of poor quality. Anthropometric information was missing for an additional 5 subjects, leaving 799 for the main analyses. The reliability of the density measurement within reading sets, given by the intraclass correlation coefficient, was excellent with estimates of 0.96, 0.99, and 0.94 for dense area, nondense area, and percent dense area, respectively. Between reading sets, the intraclass correlation coefficients were 0.98, 0.99, and 0.97 for dense area, nondense area, and percent dense area, respectively. Baseline characteristics were reported previously by Cuzick et al. (4) and are briefly reviewed in Table 1. The table also gives the means and standard deviations of each of the mammographic measures at baseline.

Figure 1 shows a scatterplot of absolute dense area by percent dense area at baseline. Percent dense area was positively correlated with dense area ($r = 0.75$), but the variation in dense area increased with increased percent dense area. Conversely, percent dense area was negatively associated with nondense area ($r = -0.80$), and the variation in nondense area decreased with increased percent dense area (Figure 2). Dense and nondense areas were also negatively correlated (graph not shown; $r = -0.36$), and the association between the computer-assisted and the visual assessments of percent dense area used in a previous paper (4) was high (graph not shown; $r = 0.85$).

Table 2 shows the linear regression coefficients (and their standard errors) for the different breast cancer risk factors for 3 of the mammographic measures. The regression estimates for percent dense area were very similar in magnitude and direction to those produced by the visual assessment reported previously (data not shown). Like the previous report (4), a multivariate model found that age, body mass index, menopausal status, predicted breast cancer risk, and smoking were all negatively associated with percent dense area. Combined, these risk factors accounted for 24% of the variation in percent dense area, with 15% of this explained by body mass index alone.

Separate examination of the components of percent dense area, dense and nondense area, revealed that, univariately, postmenopausal subjects were found to have less square-root dense area than premenopausal subjects ($\beta = -0.80$ (standard error (SE), 0.13; $P_{trend} < 0.0001$), and previous hormone therapy users had less square-root dense area than women who had never used hormone therapy ($\beta = -0.66$ (SE, 0.18); $P_{trend} = 0.002$). Neither of these covariates remained significantly associated with square-root dense area once adjusted for age in the multivariate model. In fact, only age and body mass index were significantly associated with square-root dense area in a multivariate model. Regression estimates of the association between square-root dense area
and age indicated an approximate decrease in square-root dense area of 1.5% with each yearly increase in age ($\beta = -0.070$ (SE, 0.010)). Similarly, square-root dense area decreased by approximately 0.6% with each unit increase in body mass index ($\beta = 0.029$ (SE, 0.012)). Together, age and body mass index accounted for only 7% of the variation in square-root dense area.

Regression estimates between the risk factors and square-root nondense area were opposite in direction to those of square-root dense area and both measures of percent dense area (computer-assisted and visual assessments). Postmenopausal subjects were found to have more square-root nondense area ($\beta = 0.64$ (SE, 0.14)) in the univariate analysis, but there was no evidence of this association once adjusted for age in the multivariate model. Square-root nondense area increased by approximately 0.8% with each yearly increase in age ($\beta = 0.057$ (SE, 0.009); $P_{\text{trend}} < 0.0001$) and 3.5% with each unit increase in body mass index ($\beta = 0.24$ (SE, 0.01); $P_{\text{trend}} < 0.0001$). The risk factors in the multivariate model accounted for over a third (40%) of the variation in square-root nondense area with body mass index explaining 36% alone.

Figures 3–5 are fractional polynomial smooth plots with 95% confidence intervals of all 3 mammographic measures against body mass index. Figure 3 shows a weak negative association between dense area and body mass index and an increase in the width of the confidence intervals beyond a body mass index of 30–35. Figure 4 depicts the strong negative association between percent dense area and body mass index that appears to flatten out before turning upward for body mass index values greater than 40. Like dense area,
Table 2. Linear Regression Estimates of All 3 Mammographic Measures at Entry to the Study for Various Risk Factors, International Breast Cancer Intervention Study I, 1992–2001

<table>
<thead>
<tr>
<th>Variable</th>
<th>Percent Dense Area</th>
<th>Dense Area</th>
<th>Nondense Area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Univariate</td>
<td>Multivariate</td>
<td>Univariate</td>
</tr>
<tr>
<td></td>
<td>β (SE)</td>
<td>P Value</td>
<td>β (SE)</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>−1.71 (0.15)</td>
<td>&lt;0.0001</td>
<td>−1.64 (0.14)</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>−0.52 (0.05)</td>
<td>&lt;0.0001</td>
<td>−0.0075 (0.0046)</td>
</tr>
<tr>
<td>Height, cm</td>
<td>0.31 (0.12)</td>
<td>0.01</td>
<td>0.016 (0.010)</td>
</tr>
<tr>
<td>Age at entry, years</td>
<td>−0.90 (0.12)</td>
<td>&lt;0.0001</td>
<td>−0.74 (0.15)</td>
</tr>
<tr>
<td>Age at menarche, years</td>
<td>0.49 (0.44)</td>
<td>0.3</td>
<td>−0.026 (0.036)</td>
</tr>
<tr>
<td>Age at first birth, years</td>
<td>0.2e</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nulliparious</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
</tr>
<tr>
<td>&gt;30</td>
<td>−1.76 (3.56)</td>
<td></td>
<td>0.015 (0.29)</td>
</tr>
<tr>
<td>26–30</td>
<td>−4.14 (2.61)</td>
<td></td>
<td>−0.37 (0.21)</td>
</tr>
<tr>
<td>21–25</td>
<td>−5.00 (2.41)</td>
<td></td>
<td>−0.20 (0.20)</td>
</tr>
<tr>
<td>≤20</td>
<td>−6.30 (2.75)</td>
<td></td>
<td>−0.12 (0.22)</td>
</tr>
<tr>
<td>Menopausal status at entry</td>
<td>&lt;0.0001e</td>
<td>0.02e</td>
<td>&lt;0.0001e</td>
</tr>
<tr>
<td>Premenopausal</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
</tr>
<tr>
<td>Perimenopausal</td>
<td>−1.17 (3.38)</td>
<td></td>
<td>0.22 (3.07)</td>
</tr>
<tr>
<td>Postmenopausal</td>
<td>−10.72 (1.51)</td>
<td></td>
<td>−4.80 (1.80)</td>
</tr>
<tr>
<td>Predicted risk of breast cancer</td>
<td>0.5e</td>
<td>0.03e</td>
<td>0.3e</td>
</tr>
<tr>
<td>Low (&lt;2-fold)</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
</tr>
<tr>
<td>Moderate (2–3-fold)</td>
<td>1.24 (1.60)</td>
<td>0.65 (1.42)</td>
<td>0.19 (0.13)</td>
</tr>
<tr>
<td>High (&gt;3-fold)</td>
<td>−1.41 (2.66)</td>
<td>−6.00 (2.48)</td>
<td>0.23 (0.22)</td>
</tr>
<tr>
<td>Use of hormone therapy</td>
<td>0.004e</td>
<td></td>
<td>0.002e</td>
</tr>
<tr>
<td>Never</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
</tr>
<tr>
<td>Current</td>
<td>−2.74 (1.88)</td>
<td></td>
<td>−0.18 (0.15)</td>
</tr>
<tr>
<td>Previous</td>
<td>−7.27 (2.25)</td>
<td></td>
<td>−0.66 (0.18)</td>
</tr>
<tr>
<td>Smoking status</td>
<td>0.04e</td>
<td>0.02e</td>
<td>0.15e</td>
</tr>
<tr>
<td>Never</td>
<td>Referent</td>
<td></td>
<td>Referent</td>
</tr>
<tr>
<td>Current</td>
<td>−5.05 (2.02)</td>
<td></td>
<td>−4.91 (1.79)</td>
</tr>
<tr>
<td>Previous</td>
<td>−2.15 (1.75)</td>
<td></td>
<td>−2.24 (1.56)</td>
</tr>
</tbody>
</table>

Abbreviation: SE, standard error.

a Square-root transformed.
b Adjusted for body mass index, age at entry, menopausal status at entry, predicted risk of breast cancer, and smoking.
c Adjusted for body mass index and age at entry.
d Adjusted for body mass index, age at entry, and predicted risk of breast cancer. Only significant variables were retained in the multivariate models.
e P(trend).
the variation in percent dense area increases beyond a body mass index of 30. Finally, Figure 5 shows the strong positive association between nondense area and body mass index; however, the increase in the variation of nondense area starts at body mass index values greater than 35–40.

DISCUSSION

In summary, percent dense area was negatively associated with age, body mass index, menopausal status, predicted risk, and smoking status. Nondense area was positively associated with age, body mass index, and predicted risk. Dense area was negatively associated with only age and body mass index, but to a lesser degree than for percent dense area. Although the determinants of mammographic density may not be novel, investigating how they differ for different measurements (e.g., absolute vs. percent) highlights the importance of understanding the potentially different pathways that influence both percent dense area and absolute dense area.

Although body mass index was found to be significantly associated with dense area, it explained only a fraction (<2%) of its variation in this study. Conversely, 15% of the variation in percent dense area (computer assisted or visual) and over a third of the variation in nondense area were explained by body mass index. Therefore, the negative correlation between percent dense area and body mass index appears to be largely driven by the area of nondense tissue in the breast. This implies that dense area could be a simpler predictor of breast cancer risk, as it would minimize the need to adjust for body mass index. From the results in Table 2, it appears that dense area is independent of other breast cancer risk factors as well. Assuming that dense area is also a strong predictor of risk (see below), this finding has many practical implications. It could help to simplify breast cancer risk prediction models that, in turn, could be used to optimize breast-screening intervals. It could also help to focus the search for not only environmental causes of mammographic density but also genetic causes. All 3 of the mammographic measures have been shown to have a strong heritable component (7–9). Previous work with twins showed that the genetic correlation between dense and nondense areas was significant but inverse, indicating that common genetic influences exist but likely act in opposite directions. The paper by Douglas et al. (8) suggested the exact opposite but showed significant positive genetic correlations between several measures of body size and nondense (but not dense) area. Another paper by Ursin et al. (10) suggests that a portion of the heritable variation in percent and absolute dense area may be due to variation in modifiable factors such as body mass index, and efforts to replicate this finding are underway.

There have been previous reports examining the association between breast cancer risk factors and absolute dense and nondense areas (11–15). All found a negative or no association between dense area and weight or body mass index with the exception of one study of Singaporean Chinese women (14), where body mass index was positively associated with dense area and one other study that found total body fat mass to be positively associated with dense area in premenopausal women but negatively associated in postmenopausal women (12). Those that reported associations with nondense area showed strong positive
associations with weight and/or body mass index. A recent longitudinal study by Reeves et al. (16) showed that changes in body mass index were not associated with changes in dense area but may be negatively associated with percent dense area. The authors concluded that absolute dense area is likely the better surrogate for breast cancer, but they also highlighted the importance of nondense area in the etiology of the disease as well.

The direction of the associations between the mammographic density measures and the other investigated risk factors was mostly consistent with that of the literature. A marginal negative association between percent dense area and the level of risk predicted by known risk factors was seen, but we believe this is likely to be a chance finding in view of little association seen in previous reports and a positive association with atypical hyperplasia (17). Hormone therapy use was negatively associated with absolute and percent dense areas but only in the univariate analysis, and the association did not persist after adjustment for age. We excluded the current hormone therapy users and...
repeated the regression analysis that produced the estimates in Table 2 and found no differences in the results. The women in this study were at higher risk for breast cancer and had higher than average density. We recognize that the associations of density with the risk factors may be different in different populations.

There is increasing evidence suggesting that the key parameter for breast cancer risk may be related more to the absolute amount of dense area than to the percentage of breast density (18). Several studies have investigated the association between dense area and breast cancer risk (19–25) and are summarized in Table 3. Overall, 5 found that the risk estimates were similar in magnitude to that of percent dense area, and the other 2 found that the risk of breast cancer associated with increased dense area was stronger than that of percent dense area. Maskarinec et al. (26) compared differences in percent and absolute dense areas with breast cancer risk incidence in populations at different breast cancer risk and found that age-adjusted dense area may reflect breast cancer incidence better than percent densities. None of the above studies provided risk estimates for nondense area.

The measurement of dense area was found to be as reliable as that of percent dense area, which is consistent with results from other studies (10, 26), but currently, measurement of dense area can be obtained only from digital images or digitized copies of traditional films created by using a high-resolution scanner (measurement of percent dense area is not restricted to computer-assisted methods as it can also be assessed visually). There are several sources of error when measuring mammographic density, whatever the method used. Different film types, film projections, mammography machines, and radiographer technique can all potentially alter the appearance of dense tissue in a mammogram. In particular, dense area is potentially more susceptible to variation due to the levels of compression during mammography, particularly when comparing images from the same woman over time. The level of compression may depend on radiographer technique or patient sensitivity due to menstrual phase or hormone use. In addition, there are several technical issues that must be taken into account when calculating dense area, such as traditional versus digital film, image resolution, and image size. These issues are of particular importance when comparing dense area measurements across studies.

Differences in body mass index also complicate the measurement of mammographic density, since it is a 3-dimensional object measured from a 2-dimensional image. Figures 3–5 graphically showed the associations between the mammographic measures and body mass index, which also highlighted the large variation in the measures in women with very high body mass index, particularly in dense area. This is likely due to small numbers (only 55 subjects had a body mass index greater than 35); however, it could also be related to the fact that women with higher body mass index generally have larger breasts. This could mean that either the measurement of dense area is less accurate because of measurement error in thicker breasts, or it could be that dense area is actually more variable in larger breasted women. If the first is true, the measurement error for the computer-assisted measurement of percent dense area should be more or less the same as for dense area, because it requires the calculation of the total area in addition to absolute dense area. If the second is true, then dense area is a better measure of mammographic density than percent dense area, as it captures something that percent dense area does not. A recent paper by Stuedal et al. (27) found that the association between mammographic density and breast cancer may be weaker in women with larger breasts. One possible explanation given by the authors is that breast fat could be protective due to the activity of adipocytes in fat tissue. Another explanation could be measurement error, as described above. A volumetric measurement of mammographic density could provide valuable insight into the relation between mammographic density and body mass index, as it takes into account the thickness of the breast. Several methods to measure volumetric mammographic density have been developed (28–30); however, there is currently only one study to date that has examined its association with breast cancer risk. In a recent paper by Ding et al. (31), volumetric percent density as measured by a fully automated computer program, standard mammogram form, was not found to be a better predictor of breast cancer risk than percent dense area measured via the same computer-assisted method used in this study. The authors did not report on the association between dense area and breast cancer risk nor did they adjust for body mass index.

In summary, a simple, reliable, and reproducible measure of mammographic density, independent of known breast cancer risk factors and equally predictive of risk, would be very useful for research purposes and in clinical practice. We found that absolute dense area was not associated with the multitude of risk factors that percent dense area was, thus minimizing the need to adjust for these risk factors, particularly body mass index, making it a potentially simpler biomarker for breast cancer research. It also highlights the importance of understanding the potentially different pathways that influence both percent dense area and absolute dense area. We would recommend that both percent dense area and dense area be presented whenever possible for comparisons in research.

ACKNOWLEDGMENTS

Author affiliations: Centre for Molecular, Environmental, Genetic, and Analytic Epidemiology, The University of Melbourne, Melbourne, Victoria, Australia (Jennifer Stone); Department of Radiology, Addenbrooke’s Hospital, Cambridge, United Kingdom (Ruth M. L. Warren); and Cancer Research United Kingdom Department of Epidemiology, Mathematics, and Statistics, Wolfson Institute of Preventive Medicine, Barts and The London, Queen Mary School of Medicine, University of London, London, United Kingdom (Elizabeth Pinney, Jane Warwick, Jack Cuzick).

This work was supported by Cancer Research United Kingdom program grant C569/A10404.

Conflict of interest: none declared.
REFERENCES


