Breast Density and Parenchymal Patterns as Markers of Breast Cancer Risk: A Meta-analysis

Valerie A. McCormack and Isabel dos Santos Silva
Department of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, United Kingdom

Abstract

Mammographic features are associated with breast cancer risk, but estimates of the strength of the association vary markedly between studies, and it is uncertain whether the association is modified by other risk factors. We conducted a systematic review and meta-analysis of publications on mammographic patterns in relation to breast cancer risk. Random effects models were used to combine study-specific relative risks. Aggregate data for >14,000 cases and 226,000 noncases from 42 studies were included. Associations were consistent in studies conducted in the general population but were highly heterogeneous in symptomatic populations. They were much stronger for percentage density than for Wolfe grade or Breast Imaging Reporting and Data System classification and were 20% to 30% stronger in studies of incident than of prevalent cancer. No differences were observed by age menopausal status at mammography or by ethnicity. For percentage density measured using prediagnostic mammograms, combined relative risks of incident breast cancer in the general population were 1.79 (95% confidence interval, 1.48-2.16), 2.11 (1.70-2.63), 2.92 (2.49-3.42), and 4.64 (3.64-5.91) for categories 5% to 24%, 25% to 49%, 50% to 74%, and ≥75% relative to <5%. This association remained strong after excluding cancers diagnosed in the first-year postmammography. This review explains some of the heterogeneity in associations of breast density with breast cancer risk and shows that, in well-conducted studies, this is one of the strongest risk factors for breast cancer. It also refutes the suggestion that the association is an artifact of masking bias or that it is only present in a restricted age range. (Cancer Epidemiol Biomarkers Prev 2006;15(6):1159–69)

Introduction

Breast density, a measure of the extent of radiodense fibroglandular tissue in the breast, has the potential to be used as a predictor of breast cancer risk, to monitor risk-lowering interventions and as an intermediate end point in studies of breast cancer etiology. More than 40 studies have assessed associations with Wolfe grade or percentage breast density and the majority reported 2- to 6-fold increased risks for the highest compared with the lowest risk categories; however, there has been no recent comprehensive assessment of the reasons for the wide range in estimates. Furthermore, findings from subgroup analyses (e.g., by age) have not been systematically reviewed.

Factors that may explain between-study variations in relative risk (RR) estimates include features of the study design (e.g., masking bias affects cross-sectional and cohort studies differently), population characteristics (e.g., use of general population or symptomatic controls), and methods of analysis (e.g., adjustment for confounders). There may also be important effect modifiers of this association. Differential associations by age would have wide implications for the use of this variable as an intermediate marker of risk. Some studies have found stronger associations at older (1) or younger (2) ages, whereas others have reported no age interactions (3). It is also not known whether this association differs by time since mammography, ethnicity, or any other breast cancer risk factors.

We conducted a systematic review and meta-analysis of published studies on the association between both quantitative and qualitative measures of mammographic features with the aim of exploring reasons for differences in RR estimates and, where appropriate, estimating pooled effects. The sources of heterogeneity that we explored, among others, are (a) use of concurrent or prospective mammograms, (b) use of general population or symptomatic controls/noncases, (c) age, (d) ethnicity, and (e) time since mammography.

Materials and Methods

Selection Criteria. We conducted a systematic review of all published studies investigating the association between qualitative or quantitative measures of the radiological mammographic pattern and risk of prevalent or incident breast cancer in adult women. Qualitative measures include the Wolfe grade (N1, normal fatty breast; P1 and P2, prominent ducts occupying <25% and 25-75% of the breast, respectively; and Dy, dysplastic breast with sheets of dense parenchyma) and Breast Imaging Reporting and Data System (BI-RADS) of the American College of Radiology density classification (fatty, scattered fibroglandular density, heterogeneously dense, and extremely dense). The five-grade Tabar classification was also considered (I, scollop ed contours and Cooper’s ligaments; II, evenly scattered terminal ductal lobular units; III, oval-shaped lucent areas; IV, extensive nodular and linear densities; and V, homogeneous structureless fibrosis with convex contours). Quantitative measures include percentage breast density (percentage of the mammogram with radiodense fibroglandular tissue), area of dense tissue, fractal dimension, and skewness.

Search Strategy. We identified studies that satisfied the above inclusion criteria by doing an electronic search on November 18, 2005 of Medline, EMBASE, and Pubmed databases using the following criteria: English language journals, “breast” or “mammm” in any field and “cancer,” “malign,” “carcin,” or “neopl” in the title and at least one of...
the following terms in the title: "breast densit*,” “parenchym*,” “mammo* pattern*,” “radiological pattern*,” "Wolfe*,” “Tabar*,” “BIRADS,” “mammo* feature*,” “breast pattern*,” “mammo* densit*” or "tissue densit*.” We also manually searched the references of all relevant articles to identify further studies.

**Data Extraction.** For each study, we extracted information on study design (prospective, case-control, nested case-control, or cross-sectional), study population, matching variables (if appropriate), time since mammography, use of concurrent or prospective films relative to time of diagnosis, blinding of density assessors to cancer status, age at mammography, and density classification. We recorded RRs for breast cancer associated with each breast feature and their 95% confidence intervals (95% CI)/SEs, noting the adjustments made. Where effect estimates or SEs were not reported, we calculated age-adjusted odds ratios and their 95% CIs.

**Studies of breast cancer incidence were considered to be those where exposure assessment was made with a negative mammogram; thus, cancers were diagnosed subsequently.** In prevalence studies, density was assessed in the contralateral breast of mammograms taken at diagnosis.

**Statistical Methods.** RRs of breast cancer were estimated by risk, odds, rate, or hazard ratios depending on the method of analysis. Forest plots were used to visualize study-specific RRs. Heterogeneity across studies was assessed using a $\chi^2$ statistic (Q) with inverse variance weights. The extent of inconsistency between study estimates was assessed as a percentage (I^2). Metaregression models were used to explore sources of heterogeneity. Publication bias was assessed visually from funnel plots and using the Egger test (4). We used Stata version 8.1 for analyses.

Pooled RRs were estimated using random-effects models and were only reported for similar breast density classifications.

### Table 1. Characteristics of general population studies on mammographic features in relation to breast cancer incidence

<table>
<thead>
<tr>
<th>First author, y (ref.)</th>
<th>Study design. Study population characteristics.</th>
<th>No. cases: noncases</th>
<th>Age (y)</th>
<th>Mammographic feature</th>
<th>Postmammography (y)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Byrne, 1995 (27)</td>
<td>NCC of all cases diagnosed during screening or follow-up in 22 of 29 BCDDP centers, United States, 1974-1989. Match: A, C, R, YR, no. screens. Adj: AFB, FH, EDUC, ALC, reproductive years, benign biopsy, WT.</td>
<td>1,880:2,152</td>
<td>35-74</td>
<td>Wolfe, PD</td>
<td>1-16</td>
</tr>
</tbody>
</table>

**Abbreviations:** CC, case-control; NCC, nested case-control study; A, age; AFB, age at first birth; ALC, alcohol consumption; AM, age at menarche; C, center/clinic; EDUC, education; FH, family history of breast cancer; FUP, follow-up time; HRT, current hormone replacement therapy use; HT, height; MP, menopausal status; OC, history of oral contraceptive use; P, parity; R, race/ethnicity; SEC, socioeconomic status; WT, weight; YR, year; PD, percentage density; SK, skewness; FD, fractal dimension; BCDDP, Breast Cancer Detection and Demonstration Project.

*Time from density assessment to cancer diagnosis in cases.*


Downloaded from cebp.aacrjournals.org on June 20, 2017. © 2006 American Association for Cancer Research.
Results

Included and Excluded Studies. The electronic search yielded 210 articles, 150 of which were not included: 135 were not relevant, and 15 were commentaries/reviews. Of the 60 considered studies, 3 were excluded because of unblinded exposure assessment (5) and use of a classification system, which could not be related to any others (6), and a single-twin study (7). A further 19 studies were excluded from the main analysis, as they were not independent of other included studies (studies; refs. 8-14) as well as 12 studies from Breast Cancer Detection and Demonstration Project centers (15-26) that were likely to have been subsets of two studies that included cases from all or most centers (2, 27), although they were considered in subgroup analyses. Four articles were found manually (28-31), giving a total of 42 included articles for the main analysis.

The 42 articles represent aggregate data for a total of 14,134 cases and 226,871 noncases, arising from 17 incidence studies (6,967 cases) and 17 prevalence studies (4,983 cases) in the general population and 9 studies (2,184 cases) in symptomatic populations (some studies contribute to more than one category). Incidence studies used nested case-control or cohort designs (Table 1; refs. 2, 3, 10, 27, 29, 30, 32-44), whereas prevalence studies were case-control or cross-sectional (Table 2; refs. 1, 28, 45-59). Nine studies were conducted using symptomatic populations or controls (Table 3; refs. 31, 49, 60-65).

Qualitative Classifications of Breast Density. Effect estimates from studies that analyzed Wolfe grade in four categories are shown in Fig. 1, including the study by Thomas et al. (50) that combined grades N1 and P1 as the reference category. Pooled RR estimates for both incidence and prevalence general population studies show an increasing trend with Wolfe grade with statistically significant point estimates for P2, P3, and Dyr relative to N1 of 1.76, 3.05, and 3.98 in incidence studies and lower estimates of 1.25, 1.97, and 2.42 in prevalence studies. There was no evidence of between-study heterogeneity (\( P > 0.05 \)).
Table 3. Characteristics of studies on mammographic features in relation to breast cancer risk conducted in symptomatic populations

<table>
<thead>
<tr>
<th>First author, y (ref.)</th>
<th>Study design. Study population characteristics. Matching variables if applicable (Match). Variables adjusted for in analysis (Adj)</th>
<th>No. cases: noncases</th>
<th>Age (y)</th>
<th>Mammmographic feature</th>
</tr>
</thead>
</table>
| **Studies using prospective (prediagnostic) mammograms**  
Wolfe, 1976 (62) | Cohort: mammography attendees at Hutzel hospital, United States; cases diagnosed 0.5-3.5 y after negative mammogram.  
Study 1 1967-1974 | 20,190 | 30-60+ | Wolfe |
| Study 2 1972-1975. Adj: none. | 56,528 | Wolfe |
| Egan, 1979 (60) | Cohort: referrals to Emory University Clinic, United States, 1963-1977; cases diagnosed at least 0.5 y after negative mammogram. Adj: A. | 131,634 | 35-70+ | Wolfe |
| **Studies using concurrent (diagnostic) mammograms**  
XS. Referrals for mammography at Johannesburg General Hospital, South Africa. Adj: A. | 402,1036 | 30-60+ | Wolfe |
| Brebner, 1978 (65) | XS. As for Egan 1979, but cases diagnosis and timing of breast density measurement are concurrent. Adj: A. | 300,2702 | <20-60+ | Wolfe |
| Boyd, 1982 (64) | CC. Cases: Women's College Hospital, Canada, 1973-1979; controls: diagnostic referral service. Match: A, YR. | 183,183 | 40-65 | Wolfe |
| Tabar, 1982 (49) | CC. Cases: Kopparberg County screening program, Sweden; controls: referred for mammography. Adj: A. | 183,1,674 | 40-75+ | Wolfe |

Abbreviations: CC, case-control; A, age; AFB, age at first birth; FH, family history of breast cancer; R, race/ethnicity; WT, weight; YR, year; MCC, matched case-control; XS, cross-sectional study.

excluding these two articles were higher but remained lower than that for incidence studies. Hainline et al. (45) could not be pooled, as SEs could not be estimated, but RR point estimates were consistent with overall findings (1.5, 2.7, and 4.2, respectively). Among studies conducted in symptomatic populations, RRs were very variable (heterogeneity, $P < 0.01$ for P2 and Dy relative to N1), and only one found a monotonic increasing trend (60).

Pooled RRs from studies that dichotomized the Wolfe grade (P2+Dy versus N1+P1) were highly consistent with the above findings (Fig. 2). Pooled RRs indicate that incidence and prevalence were, respectively, 1.86 and 1.44 times higher in the high-risk category. In the four studies using symptomatic controls, all found very large increased risks, but 95% CIs were wide, and there was large heterogeneity between studies.

Four studies used qualitative measures of density other than Wolfe grade. Brisson et al. reported a 2-fold increased breast cancer rate in glandular/dense compared with fatty breasts (no 95% CI was provided; ref. 2). Three studies used BIRADS classification (3, 41, 55), and, in the two that compared risks relative to the fatty parenchyma, results are consistent, giving combined RRs of 2.04 for scattered, 2.81 for heterogeneous, and 4.08 for extremely dense tissue. However, the combined estimate may be positively biased because the findings from one of these studies is restricted to postmenopausal women, as estimates were not provided for premenopausal women where “no significant associations” were found (55). Using the second category (scattered density) as the reference group, Ziv et al. also reported an increasing trend with BIRADS classification (Fig. 2; ref. 41). The single study that used Tabar classification found a 2.42-fold increased risk for pattern IV compared with pattern I but no increased risk for pattern V (Fig. 2; ref. 54).

**Quantitative Measures of Breast Density.** Studies that examined percentage density are shown in Fig. 3, including one study that assessed percentage dysplasia (28). A linear increasing trend was observed in both incidence and prevalence general population studies. Pooled RRs in incidence studies were 25% to 30% higher than in prevalence studies for categories 5% to 24% and ≥75% and were similar for categories 25% to 49% and 50% to 74%. RRs were as high as 2.92 (2.49, 3.42) and 4.64 (3.64, 5.91) for percentage density 50% to 74% and ≥75% (compared with <5%), respectively. The study by Byrne et al. contributed to at least 40% of these combined estimates (27). Despite differences in exact boundaries for breast density categories, there was no evidence of significant statistical heterogeneity (see Ps in Fig. 3).

Quantitative measures studied less frequently include the area of dense tissue and fully automated measures, such as skewness of the gray-level histogram, lacunarity, and fractal dimension. Two studies found that fractal dimension was inversely associated with risk, but the strength of the association was weaker than that for percentage density (36, 43). One study reported a negative association with regional skewness (36), but this was not replicated by another (43). Nagao et al. analyzed a fully automated measure according to whether the peak of the gray-level histogram occurred in the fat region (reference category), pectoral muscle region (RR, 2.82; 95% CI, 1.33-5.98), or had no definite peak (1.55, 0.93-2.58; ref. 57).

**Sources of Heterogeneity.** There was no evidence of publication bias in studies of percentage density and breast...
Figure 1. Association of Wolfe grades N1 (reference), P1, P2, and Dy with breast cancer risk in general (incidence and prevalence) and symptomatic populations.
cancer incidence ($Ps = 0.22, 0.94, 0.65, and 0.12, Egger test$). For prevalence studies (1, 56, 58, 59, 66), publication bias was present in the third category ($P = 0.002$), which was accounted for by the more extreme results with wide 95% CIs for premenopausal and postmenopausal women in the study by Nagata et al. (59). Exclusion of this study did not greatly alter the pooled estimate [from 2.93 (2.27-3.79) to 2.83 (2.16-3.70)].

To explore further potential sources of heterogeneity, we primarily focused on incidence studies that analyzed percentage density (10, 27, 29, 30, 37, 40, 42-44), as its risk gradient was largest, and some factors, such as time since mammography,

![Graph showing association of dichotomized Wolfe grade, BIRADS, and Tabar classifications with breast cancer risk in general (incidence and prevalence) and symptomatic populations.](image-url)

Figure 2. Association of dichotomized Wolfe grade, BIRADS, and Tabar classifications with breast cancer risk in general (incidence and prevalence) and symptomatic populations.
**Figure 3.** Association of percentage breast density with breast cancer risk in the general population (incidence and prevalence) and in symptomatic populations.

### Incidence studies, general population
(4508 cases : 8342 non-cases)

<table>
<thead>
<tr>
<th>Percentage Breast Density</th>
<th>Combined Relative Risk (95% CI)</th>
<th>Heterogeneity p-value, I^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-24% v &lt;5%</td>
<td>1.79 (1.48, 2.16)</td>
<td>p=0.22, I^2=27%</td>
</tr>
<tr>
<td>Boyd 1995 (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Byrne 1995 (27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maskarinec 2000 (29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Gils 2000 (40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ciato 2004 (44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maskarinec 2005 (42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torres-Mejia 2005 (43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-49% v &lt;5%</td>
<td>2.11 (1.70, 2.63)</td>
<td>p=0.09, I^2=46%</td>
</tr>
<tr>
<td>Boyd 1995 (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Byrne 1995 (27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kato 1995 post (37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maskarinec 2000 (29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomas 2002 (30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maskarinec 2005 (42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torres-Mejia 2005 (43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-74% v &lt;5%</td>
<td>2.92 (2.49, 3.42)</td>
<td>p=0.63, I^2=0%</td>
</tr>
<tr>
<td>Boyd 1995 (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Byrne 1995 (27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kato 1995 pre (37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maskarinec 2000 (29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Gils 2000 (40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomas 2002 (30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ciato 2004 (44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maskarinec 2005 (42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Torres-Mejia 2005 (43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75%+ v &lt;5%</td>
<td>4.64 (3.64, 5.91)</td>
<td>p=0.50, I^2=0%</td>
</tr>
<tr>
<td>Boyd 1995 (10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Byrne 1995 (27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thomas 2002 (30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ciato 2004 (44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Prevalence studies, general population
(2219 cases : 4063 non-cases)

<table>
<thead>
<tr>
<th>Percentage Breast Density</th>
<th>Combined Relative Risk (95% CI)</th>
<th>Heterogeneity p-value, I^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-24% v &lt;5%</td>
<td>1.39 (1.10, 1.76)</td>
<td>p=0.90, I^2=0%</td>
</tr>
<tr>
<td>Boyd 1992 (28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brisson 2003 (56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ursin 2003 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kerlikowske 2005 (68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagata 2005 post (59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagata 2005 pre (59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-49% v &lt;5%</td>
<td>2.22 (1.75, 2.81)</td>
<td>p=0.90, I^2=0%</td>
</tr>
<tr>
<td>Boyd 1992 (28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brisson 2003 (56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ursin 2003 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kerlikowske 2005 (68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagata 2005 post (59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagata 2005 pre (59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-74% v &lt;5%</td>
<td>2.93 (2.27, 3.79)</td>
<td>p=0.91, I^2=0%</td>
</tr>
<tr>
<td>Boyd 1992 (28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brisson 2003 (56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ursin 2003 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kerlikowske 2005 (68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagata 2005 pre (59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75%+ v &lt;5%</td>
<td>3.67 (2.72, 4.96)</td>
<td>p=0.51, I^2=0%</td>
</tr>
<tr>
<td>Boyd 1992 (28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brisson 2003 (56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ursin 2003 (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kerlikowske 2005 (68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagata 2005 pre (59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Symptomatic populations
Wolfe 1997 (31) (170 cases: 170 controls)

<table>
<thead>
<tr>
<th>Percentage Breast Density</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-49% v &lt;5%</td>
<td>Decreased risk</td>
</tr>
<tr>
<td>50-74% v &lt;5%</td>
<td>Increased risk</td>
</tr>
<tr>
<td>75%+ v &lt;5%</td>
<td>NA</td>
</tr>
</tbody>
</table>

could not be examined in prevalence studies. The effect of masking bias was evident, as a lower RR was observed in the highest density category among studies that excluded cancers that arose during the first year (10, 27, 30, 40) after a negative screen compared with those that included them (Fig. 4; refs. 29, 37, 42-44). This difference was largely influenced by the study of Ciatto et al. (44), in which all cases were diagnosed within 2 years of screening (as this study was designed to investigate associations with interval cancers rather than with overall breast cancer risk). This study was removed from further analyses.

We assessed associations with short- (10, 27, 29, 30, 37, 42) and long-term (27, 30, 40, 43) risk by separating studies according to whether the majority of mammograms were taken within or >5 years previously. Two studies contributed to both groups, as they presented results by time since mammography (27, 30). A stronger trend was observed for short-term risk than for long-term risk, especially for the highest density category where the test for heterogeneity was borderline significant (Fig. 4; \( P = 0.06 \)). This difference remained when we restricted analyses to studies that excluded cancers in the first year of follow-up. These results largely reflect the weakening, with time since mammography, of the association in the study by Thomas et al. (30), but the larger study by Byrne et al. (27) did not observe this trend.

Pooled estimates from four incident studies of percentage density where it was possible to analyze younger (premenopausal or <50 years at the time of mammography) and older women separately are shown in Fig. 4 (10, 27, 30, 37). There are some suggestions of stronger RRs at older ages for older women separately are shown in Fig. 4 (10, 27, 30, 37). Parity or number of deliveries has been reported as an effect modifier in two studies, with one finding a stronger association of percent-

density in nulliparous women, whereas the other found that women with Tabar grades IV and V were at an increased risk only if they had one or two deliveries but not more (13, 40). Herrinton et al. found no evidence of effect modification by alcohol intake (26), and two studies found no differences according to ever use of hormone replacement therapy (13, 55). Boyd et al. found that associations of percentage density with breast cancer risk were stronger in women with a first-degree relative with breast cancer (9).

Discussion

The combined data presented here confirm that breast density, measured using either Wolfe grade or percentage density, is strongly associated with breast cancer risk, as determined by general population studies of either incident or prevalent cancer risk. Pooled RRs were higher for percentage density than for Wolfe grade. Individual studies that were able to assess the two measures simultaneously have suggested that Wolfe grade is not predictive after adjusting for percentage density (56). However, the breast density-breast cancer association only held in general population studies. The inconsistent findings from studies conducted within symptomatic populations may reflect differing extents of underlying breast disease. Breast density is also a risk factor for benign breast disease, so RRs were expected to be lower.

In general population studies, there was evidence that masking bias affected study results as estimates were, as expected, consistently lower in prevalence than in incidence studies and lower in incidence studies after excluding cancers diagnosed in the first-year postmammography. In prevalence studies, in which breast density measurement and diagnosis were conducted using mammograms taken concurrently, RRs are expected to be underestimated, as there is lower mammographic sensitivity and, hence, more false negatives in dense than in fatty breasts. On the contrary, during the follow-up of a screen-negative cohort in a study of breast cancer incidence, excess cancers in dense breasts may actually be prevalent cancers that were initially undetected, leading to an overestimation of RRs.

Strong linear trends were observed for percentage density, despite RRs referring to different (but overlapping) ranges. RR estimates were consistent within each density category, suggesting that this method, although not ideal, was reasonable in this situation. The combined pooled RRs for categories 5% to 24%, 25% to 49%, 50% to 74%, and ≥75% relative to <5% were 1.79 (1.48-2.16), 2.11 (1.70-2.63), 2.92 (2.49-3.42), and 4.64 (3.64-5.91), respectively, for incidence studies. The true association may be even stronger, as nondifferential measurement error of breast density would lead to underestimation of associations. Assuming moderate error (intraclass correlation of 0.90; ref. 27), RR estimates for categories 50% to 74% and ≥75%, when corrected, would increase to 3.29 and 5.50, respectively. New automated methods for density...
measurement based on estimation of the volume of dense tissue are currently being developed. Demonstration that their RRs are as strong as those reported here will be needed if they are to serve as useful alternatives to current methods of breast density measurement.

Other classifications and measurements of breast density have been assessed less frequently. Studies of the BIRADS classification suggest that this four-category grade is also a strong predictor of breast cancer risk, although the magnitude may be slightly lower than that for percentage density. However, as this density grade is routinely recorded at the time of mammography in the United States (67), the BIRADS classification could be used in epidemiologic studies within this setting. Other measures, such as skewness and fractal dimension, have only been examined in two studies whose findings were not entirely consistent, and their effect estimates were not as strong as those for percentage density (36, 43). Debate still exists about whether

**Figure 4.** Combined RR estimates for the association of percentage breast density with incident breast cancer risk (reference category <5%) by potential sources of heterogeneity.

Combined relative risks for breast cancer associated with categories of percent breast density, in subgroups of incidence studies.

*Total number of cases: noncases in contributing studies. Contributing studies are: Boyd 1995 (10), Byrne 1995 (27), Maskarinec 2000 (29), Thomas 2002 (30), Kato (37), van Gils 2000 (40), Maskarinec 2005 (42), Torres-Mejia 2005 (43), Ciatio 2004 (44).
Breast Density and Breast Cancer Risk: A Meta-analysis

percentage density or area of dense tissue is the strongest predictor of risk (43).

The increased risk associated with breast density is independent of other known risk factors for this disease but is confounded by age and BMI. Failure to take BMI into account in postmenopausal women explained some heterogeneity, as adjustment for this confounder leads to a strengthening of the association. We found little evidence of interactions between other risk factors for breast cancer and breast density. Importantly, the combined data suggest that breast density measured at both premenopausal and postmenopausal ages is a marker of subsequent breast cancer risk and that there is no clear evidence that the strength of this association differs between these ages. However, most premenopausal women in these studies were in their late 30s and 40s, so it is still not known whether mammographic density in early adulthood is predictive of risk in later life. From the available data, there is no evidence to suggest that this association does not hold in all studied ethnic groups (Caucasian, Asian, and African American). Future studies based on or including women of other ethnicities would be informative.

Two of the included studies examined changes in the strength of the association with time since mammography, but only one suggested that the association weakened over time (27, 30). Further investigation of the duration for which an increased RR is applicable is needed. Breast density measured >5 years previously will not capture more recent changes in density, which may influence risk. This problem may be more acute at menopausal ages when density reduces rapidly.

The strength of the association of breast density with breast cancer risk is greater than that for most other established breast cancer risk factors, with the exception of age and some genetic factors. Breast density also accounts for a large number of cases. Assuming that ~20% and 10% of premenopausal women have densities of 50% to 74% and ≥75%, respectively, the population attributable fraction is 26.7% for densities over 75% and 42.8% for densities over 50%. At postmenopausal ages, the corresponding fractions would be 9.8% and 23.2% (based on 10% and 3% with densities of 50-74% and ≥75%). However, there was no threshold level below which density was not associated with risk. Shifting the entire breast density distribution downwards by a few percentage (if possible) might reduce overall breast cancer rates. Breast density may be amenable to change. A comprehensive review of correlates of breast density and the importance of this marker has great potential to be used for research into the etiology and prevention of breast cancer.

References

Breast Density and Parenchymal Patterns as Markers of Breast Cancer Risk: A Meta-analysis

Valerie A. McCormack and Isabel dos Santos Silva


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/15/6/1159

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2006/06/15/15.6.1159.DC1

Cited articles
This article cites 69 articles, 24 of which you can access for free at:
http://cebp.aacrjournals.org/content/15/6/1159.full#ref-list-1

Citing articles
This article has been cited by 100 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/15/6/1159.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, contact the AACR Publications Department at permissions@aacr.org.