Cancer Epidemiology, Biomarkers & Prevention

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: doi:10.1158/1055-9965.EPI-06-0034
Supplementary Material	Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2006/06/15/15.6.1159.DC1.html

Cited Articles	This article cites 69 articles, 37 of which you can access for free at: http://cebp.aacrjournals.org/content/15/6/1159.full.html#ref-list-1
Citing Articles	This article has been cited by 68 HighWire-hosted articles. Access the articles at: http://cebp.aacrjournals.org/content/15/6/1159.full.html#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.

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 $\geq$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 (N_1, normal fatty breast; P_1 and P_2, 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 (BIRADS) 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, scalloped 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 ''mamm*'' in any field and ''cancer*,'' ''malig*,'' ''carcin*,'' or ''neopl*'' in the title and at least one of

the following terms in the title: ''breast densit*,'' ''paren-chym*,'' ''mammo* pattern*,'' ''radiological pattern*,'' ''Wolfe*,'' ''Tabar*,'' ''BIRAD*,'' ''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 χ^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

First author, y (ref.)	Study design. Study population characteristics. Matching variables if applicable (Match). Variables adjusted for in analysis (Adj)	No. cases: noncases	Age (y)	Mammographic feature	Postmammography (y)*
Verbeek, 1984 (32)	CC. Cases: diagnosed at 4th screening round in 1981-1982, Nijmegen screening program, the Netherlands; controls: 4 screening rounds. Adj: A, AM, P, AFB, OC, MP, BMI.	20:60	35-64	Wolfe	5-6
Brisson, 1988 (2)	Cohort of white women in follow-up of BCDDP, United States, 1974-1985. Adj: A.	1,999:53,054	35-74	3 grades	1-9
de Stavola, 1990 (33)	Cohort: Guernsey prospective study, United Kingdom. 1977-1988. Adj: A, FUP, BMI.	58:4,896	30-75+	Wolfe	0.5 to ~7
Ciatto, 1993 (34)	Cohort: Florence district population screening, Italy. 1985-1991. Adj: A.	126:17,745	40-70	Wolfe	>0 to 5
Boyd, 1995 (10); Byng, 1998 (35); Yaffe, 1998 (36)	NCC. Canadian National Breast Screening Study. Match: YR, A, FUP, C. Adj: AM, P, AFB, FH, HT, WT.	354:354	40-59	PD, SK, FD	1-5
Byrne, 1995 (27)	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.	1,880:2,152	35-74	Wolfe, PD	1-16
Kato, 1995 (37)	NCC. New York University Women's Health Study, United States, 1985-1991. Match: A, MP, YR, no. blood samples. Adj: P, BMI, time since menopause.	52:195 (pre) 91:178 (post)	35-65	Wolfe, PD	0.5-6
Thurfjell, 1996 (38)	NCC. Uppsala county screening program, Sweden. 1988-1993. Match: A. Adj: A, breast size.	295:589	40-74	Wolfe	1-5
Salminen, 1998 (39)	Cohort of pilot screening program, Kofta, Finland. 1982-1993. Adj: A.	68:3,994	40-47	Wolfe	0.5-10
Maskarinec, 2000 (29)	NCC. Kaiser Permenente, Hawaii, United States, 1991-1997. Match: YR, A, R. Adj: AM, MP, P, AFB, FH, HRT, previous breast problems.	647:647	Mean 60	PD	Mean 0.5
van Gils, 2000 (40)	NCC. Setting same as for Verbeek 1984, cases diagnosed, 1985-1994. Match: YOB, YR. Adj: P.	129:517	36-72	PD	10-11
Thomas, 2002 (30)	CC. 4 previous CC studies combined, Seattle, United States, 1983-1995. Match (frequency): A. Adj: A, C.	547:472	<50	Wolfe, PD	1 to NK
Ciatto, 2004 (44)	NCC of interval cancers. Florence district, Italy. 1996-1999. Match: A, YR. Adj: A. no. breast views, screening round.	90:360	≥50	PD	0 to <2
Vacek, 2004 (3)	Cohort in Vermont Breast Cancer Surveillance System, United States, 1997-2001. Adj: FH, AFB, HRT, BMI.	1,191:60,653	35-75+	BIRADS	1 to ~3.1
Ziv, 2004 (41)	Cohort of women who had mammography between 1995-2002 at San Francisco Mammography Registry. Adj: A, BMI, HRT, FH, MP, P, R.	701:44,110	Mean 58 in cases	BIRADS	0-7.5
Maskarinec, 2005 (42)	NCC. Multiethnic cohort study, Hawaii, United States, 1993-2000. Match (frequency): A, R. Adj: R, A, BMI, AFB, P, AM, age at menopause, HRT, FH.	607:667	Mean 60	PD	0-7
Torres-Mejia, 2005 (43)	Cohort: Guernsey prospective study, United Kingdom. 1986-2003. Adj: A, EDUC, SEC, job, P, HT, BMI, recent change in BMI.	111:3,100	35-80	Wolfe, PD, FD, SK	0-17

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.

Table 2. Characteristics of general population studies on mammographic features in relation to breast cancer prevalence

First author, y (ref.)	Study design. Study population characteristics. Matching variables if applicable (Match). Variables adjusted for in analysis (Adj)	No. cases: noncases	Age (y)	Mammographic feature
Hainline, 1978 (45)	MCC. Cases: Duke Hospital, United States, 1974-1976; controls: asymptomatic screenees. Match: A, R.	171:171	35-70+	Wolfe*
Weich, 1981 (46)	CC. Cases: Rambam Medical Centre; controls: screening program, Israel. 1974-1979. Adj: A.	442:1,000	30-60+	Wolfe
Boyd, 1982 (28)	MCC. Cases: Women's College Hospital, Canada, 1973-1979; controls: screening volunteers. Match: A, YR.	183:183	36-65	Wolfe, PD
Brisson, 1982 (47)	CC. Case: 2 U.S. hospitals, 1972-78; controls: referred for routine/baseline mammogram. Adj: none.	408:1,021	20-60+	Wolfe, PN, PH[†]
Janzon, 1982 (48)	XS. Malmo breast cancer screening project, Sweden. 1977-NK. Adj: A.	118:14,992	40-69	Wolfe
Tabar, 1982 (49)	XS. Consecutive screens at Kopparberg County screening programme, Sweden. Adj: A.	187:26,970	40-75+	Wolfe
Chaudary, 1983 (50)	CC. Cases: Guy's hospital London 1980-1982; controls: Guernsey screening, United Kingdom. Adj: none.	60:937	NK	Wolfe
Brisson, 1984 (51)	CC. Cases: 3 hospitals in Boston and New Jersey, United States, 1978-1979; controls: referred for a ''routine'' mammogram. Adj: A, HT, WT.	362:686	Pre, post	Wolfe, PN, PH[†]
Horwitz, 1984 (52)	MCC. Cases: Yale-New Haven Hospital, United States, 1976-1979; controls: women with clinically normal mammograms. Match: A, R. Adj: no other.	105:103	>45 and post	Wolfe
Brisson, 1989 (53)	CC. Cases: treated in Quebec, 1982-1984; controls: Quebec screening study, Canada. Adj: A, WT, P, EDUC.	290:645	40-62	Wolfe, PD[†]
Jakes, 2000 (54)	MCC. Cases: prevalent and incident screen cancers, Singapore, 1995-1996; controls: negative screen cancer. Match: A. Adj: AM, AFB, breast feeding.	174:348	45-69	Tabar
Lam, 2000 (55)	MCC. Vermont breast cancer surveillance system, United States, 1996-1997. Match: A, WT.	381:1,547	Post	BIRADS
Brisson, 2003 (56)	CC. Pooling of Brisson, 1982 (47), 1984 (51), and 1989 (53). Adj: A, WT.	1,060:2,352	20-95	PD
Nagao, 2003 (57)	MCC. Cases: Gifu hospital, Japan, 1998-1999; controls: screening. Match: A, number of deliveries. Adj: AFB, no. births.	237:742	NK	Wolfe, other
Ursin, 2003 (1)	CC. Two CC studies at University of Southern California, United States, 1994-1998. Match (frequency): A, R. Adj: A, BMI, P, HRT, AM, FH, MP, AFB.	622:433	35-64	PD
Kerlikowske, 2005 (58)	CC. San Francisco Mammography Register, United States, 1992-2002. Adj: A, FH, AFB, R, BMI, bone mineral density.	208:436	Mean 63.6 (SD, 9.4)	PD
Nagata, 2005 (59)	CC. Cases and controls: Gifu City Hospital mass screening program, Japan, 2000-2002. Adj: A, BMI, AM, AFB, P, HRT, breast feeding, FH, MP.	71:370 pre, 75:289 post	30-80	PD

Abbreviations: MCC, matched case-control; XS, cross-sectional study; PN, percentage nodularities; PH, percentage homogeneous densities; post: postmenopausal; pre, premenopausal.
*Hainline, 1978 does not appear in Figs., as no estimates of precision were provided.
[†] Results for percentage nodularity do not appear in Figs., as these three studies are subsets of Brisson, 2003, which is included.

and for independent subjects. Tests of heterogeneity were based on fixed-effects models to maximize power. For percentage density, we combined RR estimates using the five most commonly used categories of ~ <5% (reference), 5% to 24%, 25% to 49%, 50% to 74%, and ≥75%. Where categorizations from individual studies differed, we placed each RR into the category where its midpoint fell.

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. (30) that combined grades N_1 and P_1 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 P_1, P_2, and Dy relative to N_1 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 ($P > 0.05$), with the exception of the RRs for P_2 relative to N_1 in prevalence studies, where the studies of Weich et al. (46) and Nagao et al. (57) were clear outliers (Fig. 1). Combined RRs

Table 3. Characteristics of studies on mammographic features in relation to breast cancer risk conducted in symptomatic populations

First author, y (ref.)	Study design. Study population characteristics. Matching variables if applicable (Match). Variables adjusted for in analysis (Adj)	No. cases: noncases	Age (y)	Mammographic feature
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.		30-60+	Wolfe
Study 1	1967-1974	20:1,910		
Study 2	1972-1975. Adj: none.	56:5,228		
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.	131:6,334	35-70+	Wolfe
Studies using concurrent (diagnostic) mammograms				
Peyster, 1977 (63)	CC. Cases and controls: referrals to Massachusetts General Hospital, United States, 1972-1976. Adj: A.	402:1,036	30-60+	Wolfe
Brebner, 1978 (65)	XS. Referrals for mammography at Johannesburg General Hospital, South Africa. Adj: A.	300:2,702	<20-60+	Wolfe
Egan, 1979 (60)	XS. As for Egan 1979, but cases diagnosis and timing of breast density measurement are concurrent. Adj: A.	658:13,465	30-70+	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
Kojima, 1984 (61)	XS. Women who underwent xeromammography at Kyoto Prefectual University of Medicine, Japan, 1978-1982. Majority had clinical breast disease. Adj: A.	91:746	30-60+	Wolfe
Wolfe, 1987 (31)	MCC. Cases: Hutzel hospital, United States, 1979-1982; controls: screened at the same hospital, generally symptomatic. Match: YR, A, R, AFB, FH. Adj: no other.	160:160	30-85	Wolfe, PD

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 P_2 and Dy relative to N_1), and only one found a monotonic increasing trend (60).

Pooled RRs from studies that dichotomized the Wolfe grade (P_2+Dy versus N_1+P_1) 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.
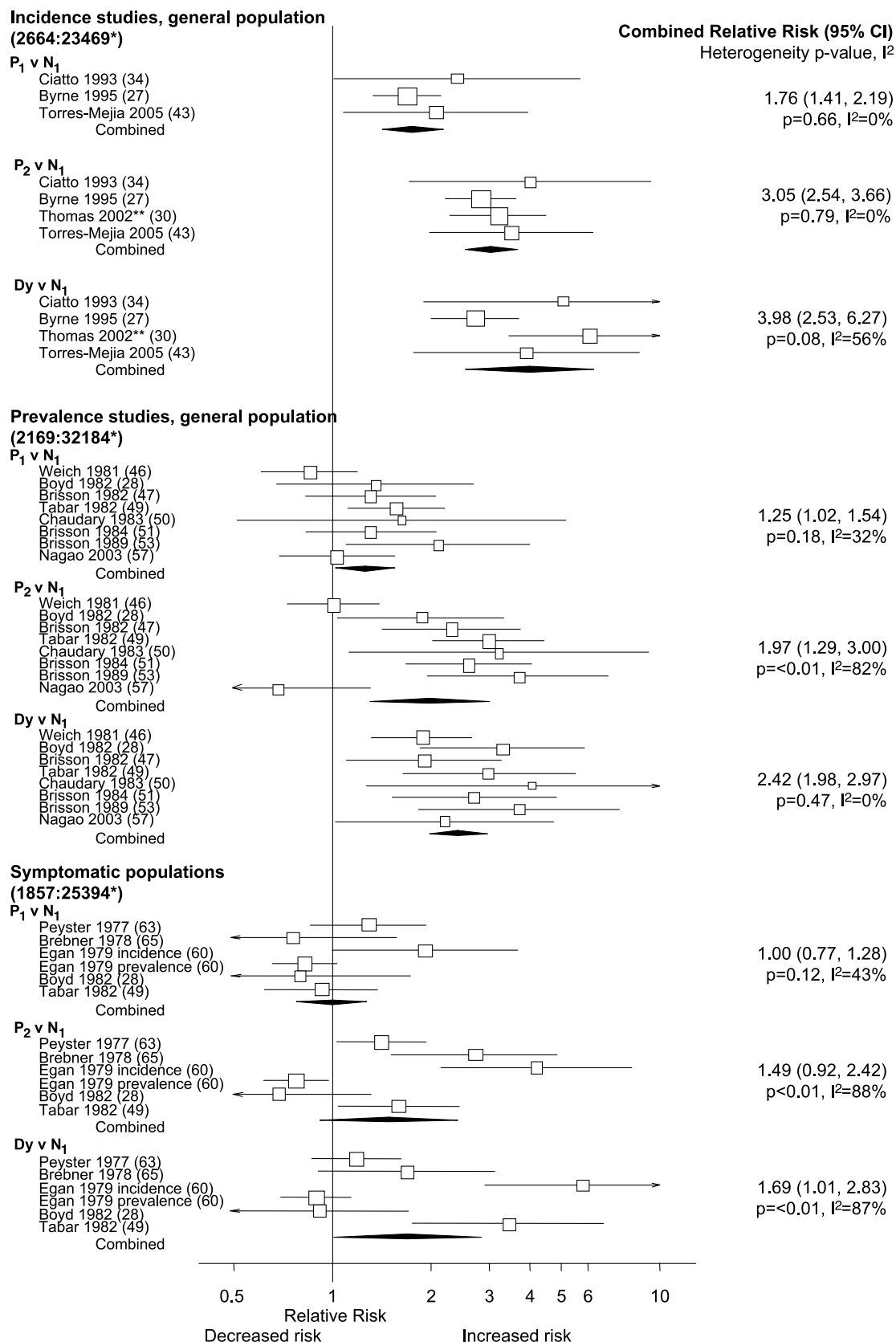
Four studies used qualitative measures of density other than Wolfe grade. Brisson et al. reported ~2-fold increased breast cancer rate in glandular/dense compared with fatty breasts (no 95% CI was provided; ref. 2). Three studies used BIRADS classification (3, 41, 55), and, in the two that compared risks 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 N_1 (reference), P_1, P_2, and Dy with breast cancer risk in general (incidence and prevalence) and symptomatic populations.
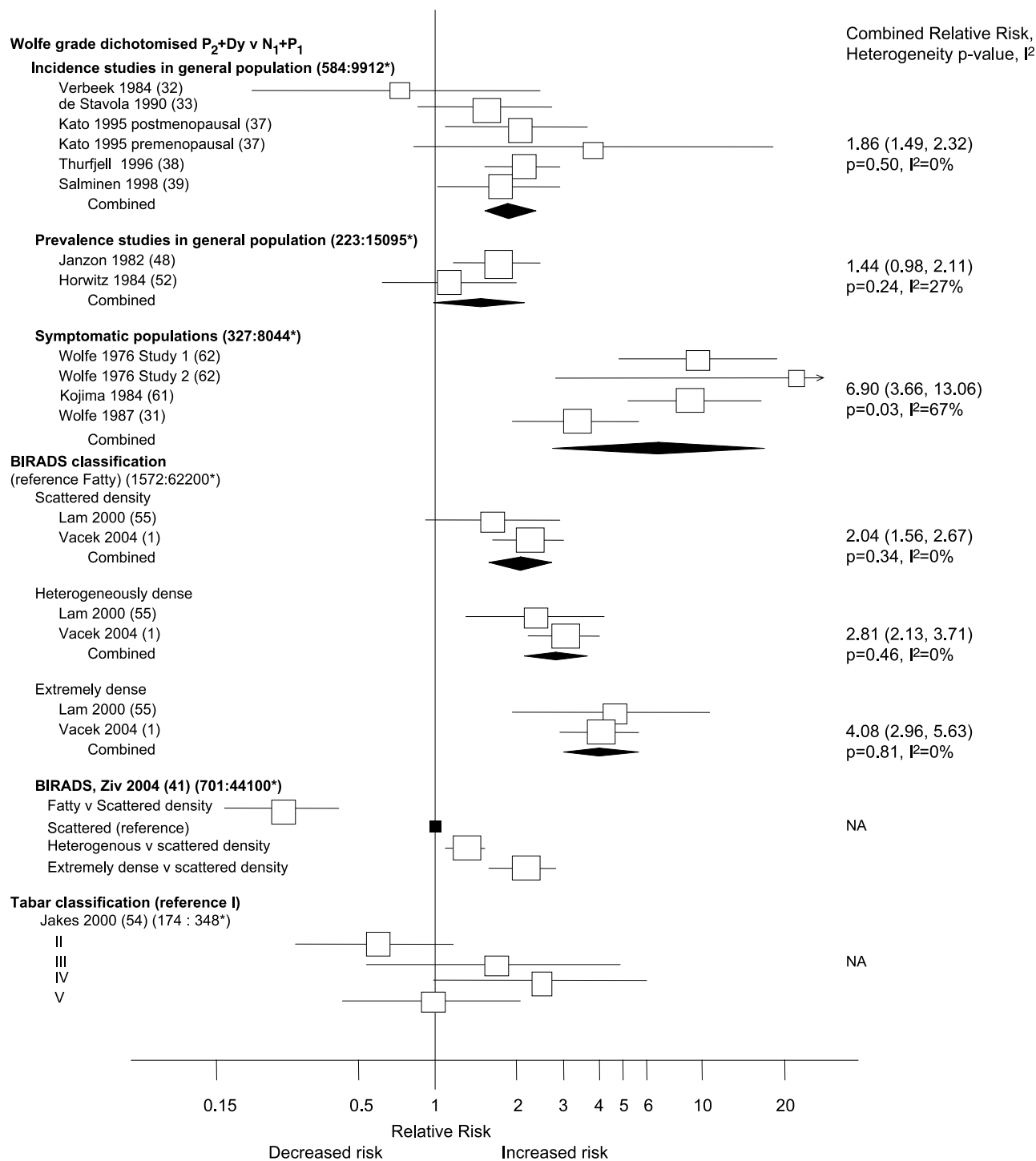
Figure 2. Association of dichotomized Wolfe grade, BIRADS, and Tabar classifications with breast cancer risk in general (incidence and prevalence) and symptomatic populations.

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,

Figure 3. Association of percentage breast density with breast cancer risk in the general population (incidence and prevalence) and in symptomatic populations.

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 category 25% to 49% relative to <5% ($P = 0.05$), but differences between age groups for other categories were not significant and not consistent. There is no suggestion that this association is restricted to either younger or older women. On examination of the large prevalent studies that presented data by menopausal status or age, Ursin et al. (1) also found a significantly stronger association at an age of >50 years ($P = 0.05$). However, two other large studies had contrary findings. Brisson et al. (2) found a stronger association in younger, whereas Vacek et al. (3) found no significant difference in the relationship at premenopausal and postmenopausal ages for the association with BIRADS classification.

Although the majority of studies were conducted in Caucasian populations, four were carried out in Asia, and their findings were consistent with the overall pooled effects (54, 57, 59, 61). Four U.S. studies presented ethnic group-specific associations (1, 29, 31, 42). Two of them commented on associations in African American women, but neither found statistically significant differences: Ursin et al. (1) found a slightly weaker association in African American women (RR, 1.66; 95% CI, 0.64-4.32) compared with that in White women (RR, 2.56; 95% CI, 1.23-5.31 for ≥60% breast density compared with <10%), whereas Wolfe et al. (31), in a study of symptomatic women, found a slightly stronger association in the former group. In two Hawaiian studies, weaker associations were observed in Asian (Japanese) women than in Caucasian women (29, 42). Duffy et al. also reported modifications by ethnic group in a case-control study in Singapore ($P = 0.05$), as they found that non-Chinese women with dense patterns were at an increased risk but that Chinese women were not (13). On the contrary, Ursin et al. found a larger association in Asian Americans than in White or African Americans (odds ratios, 1.30, 1.15, and 1.10, respectively, for a 10% increase in percentage density; ref. 1).

Within studies, authors have reported age and body size [weight or body mass index (BMI)] to be negative confounders. Among incidence studies of percentage density, age was controlled for in all nine studies. In studies that controlled for body size (10, 27, 37, 42, 43), pooled RRs were 1.78, 2.46, 3.02, and 4.59, respectively, for categories 5% to 24%, 25% to 49%, 50% to 74%, and ≥75% relative to <5%, which were generally larger than the corresponding pooled RRs for studies that did not make this adjustment (1.81, 1.56, 2.67, and 4.40; refs. 29, 40, 43). Two studies (using Tabar classification and percentage density) reported stronger associations in women with higher BMIs, but the test for heterogeneity was not significant in either study (13, 42).

Other potential effect modifiers have been examined only by a small minority of studies, so it was not possible to produce subgroup-specific RRs; instead, we simply quote the findings reported in individual studies. Parity or number of deliveries has been reported as an effect modifier in two studies, with one finding a stronger association of percentage 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
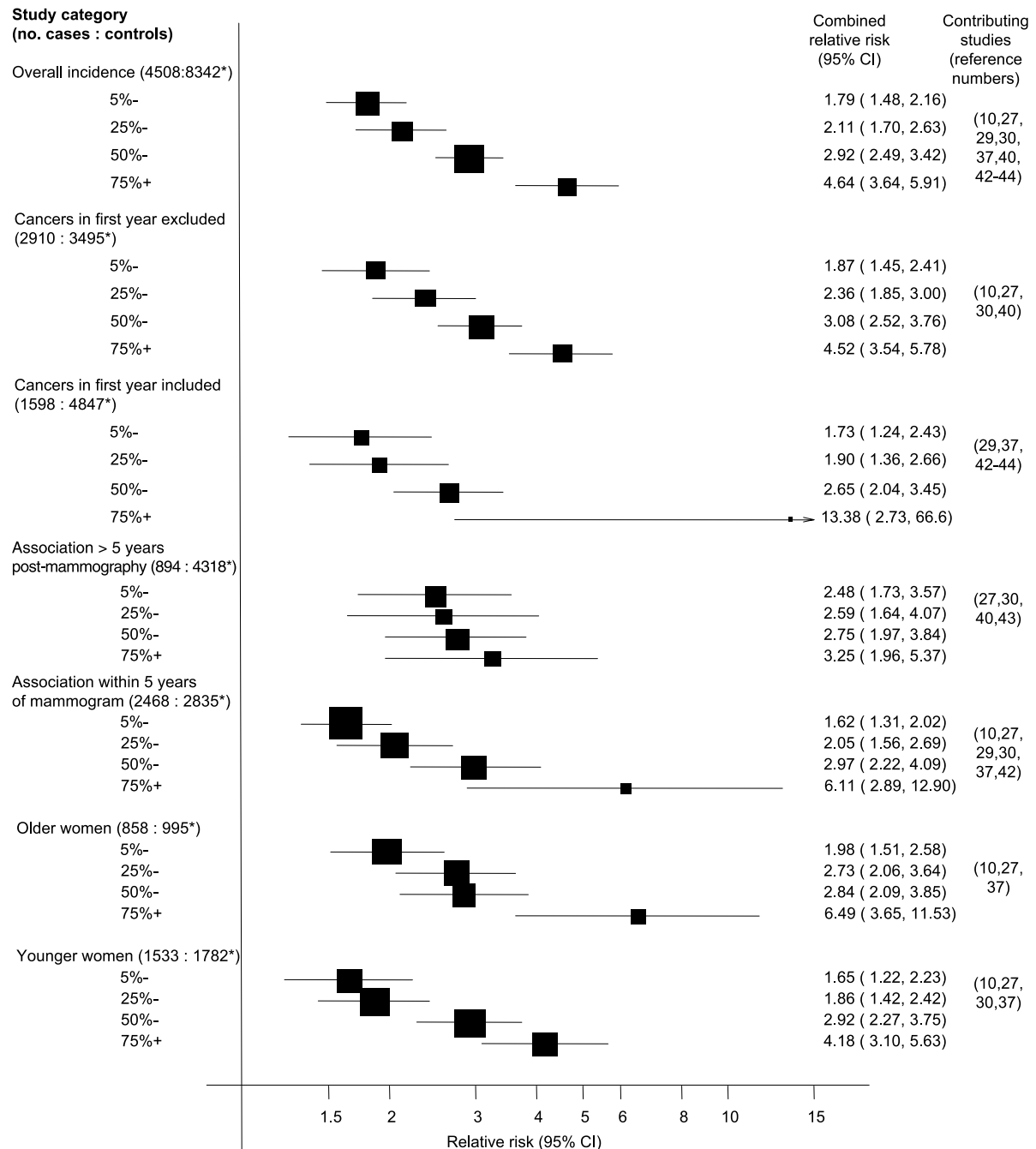
Combined relative risks for breast cancer associated with categories of percent breast density, in subgroups of incidence studies.
*Total number of cases : noncases in contributing studies. Contributing studies are: Boyd 1995 (10), Byrne 1995 (27), Maskarinec 2000 (29), Thomas 2002 (30), Kato (37), van Gils 2000 (40), Maskarinec 2005 (42), Torres-Mejia 2005 (43), Ciatto 2004 (44).

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.

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 magni-tude 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 skew-ness 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 per-centage density (36, 43). Debate still exists about whether

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 recently been conducted (68). Although it has a strong genetic component (69), it is also influenced by lifestyle factors, such as diet (70, 71), and it can be reduced by drugs, such as tamoxifen (72). Controversy remains, however, over whether women should be informed of their density, as it is not known whether reductions in breast density would necessarily lead to lower breast cancer risk. In terms of screening policy, women with dense breasts may benefit from shorter screening intervals. For research purposes, breast density may be used as an intermediate or additional end point in studies of breast cancer. Candidate susceptibility genes for breast cancer may be initially identified using breast density as an intermediate phenotype.

The mechanisms through which breast density affects breast cancer risk are not fully understood (68). Density reflects the proportion of epithelial and stromal tissue in the breast as opposed to nondense fatty tissue. Breast cancers originate in epithelial cells, so greater areas of fibroglandular tissue may reflect a greater number of cells that are at risk of carcinogenesis and/or an increased rate of epithelial proliferation. It is plausible that many of the established breast cancer risk factors influence risk through their effect on density. There is striking similarity between the breast density-age profile and the model for the rate of breast tissue ageing as proposed by

Pike et al. to explain the incidence of breast cancer with age (73), and, corresponding to this model, cumulative exposure to breast density would be the pertinent risk factor for which density at a given age is a good proxy.

In summary, some of the heterogeneity in associations of mammographic density with breast cancer risk can be explained by differences in density classification, use of concurrent or prospective mammograms, and inclusion of general or symptomatic populations. RRs 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 BIRADS 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. Well-conducted incidence studies suggest that increasing breast density is associated with an increased risk of breast cancer and that the magnitude of this association is 4.64-fold (3.64-5.91) for the most dense (≥75%) compared with the least dense category (<5%). Given the large population attributable risk of mammographic density and that mammography is conducted regularly for screening purposes, the routine measurement of mammographic density should be given more consideration. This marker has great potential to be used for research into the etiology and prevention of breast cancer.

References

1. 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.
2. Brisson J, Morrison AS, Khalid N. Mammographic parenchymal features and breast cancer in the breast cancer detection demonstration project. J Natl Cancer Inst 1988;80:1534–40.
3. Vacek PM, Geller BM. A prospective study of breast cancer risk using routine mammographic breast density measurements. Cancer Epidemiol Biomarkers Prev 2004;13:715–22.
4. Sutton AJ. Methods for meta-analysis in medical research. Wiley series in probability and statistics. New York: Wiley; 2000.
5. Mendell L, Rosenbloom M, Naimark A. Are breast patterns a risk index for breast cancer? A reappraisal. AJR Am J Roentgenol 1977;128:547.
6. Xu K, Ao-rong T, Bing-zhang S, et al. Relationship between breast parenchymal pattern and breast cancer incidence. Chin Med J (Engl) 1986;99:277–80.
7. Kaprio J, Alanko A, Kivisaari L, Standertskjold-Nordenstam C. Mammographic patterns in twin pairs discordant for breast cancer. Br J Radiol 1987;60:459–62.
8. Brisson J. Family history of breast cancer, mammographic features of breast tissue, and breast cancer risk. Epidemiology 1991;2:440–4.
9. Boyd NF, Lockwood GA, Martin LJ, et al. Mammographic densities and risk of breast cancer among subjects with a family history of this disease. J Natl Cancer Inst 1999;91:1404–8.
10. 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.
11. Egan RL, Mosteller RC. Breast cancer mammography patterns. Cancer 1977;40:2087–90.
12. Gravelle I, Bulstrode J, Bulbrook R, et al. A prospective study of mammographic parenchymal patterns and risk of breast cancer. Br J Radiol 1986;59:487–91.
13. Duffy SW, Jakes RW, Ng FC, Gao F. Interaction of dense breast patterns with other breast cancer risk factors in a case-control study. Br J Cancer 2004;91:233–6.
14. Wolfe J. Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer 1976;37:2486–92.
15. Wilkinson E, Clopton C, Gordonson J, et al. Mammographic parenchymal pattern and the risk of breast cancer. J Natl Cancer Inst 1977;59:1397–400.
16. Krook PM. Mammographic parenchymal patterns as risk indicators for incident cancer in a screening program: an extended analysis. AJR Am J Roentgenol 1978;131:1031–5.
17. Moskowitz M, Gartside P, McLaughlin C. Mammographic patterns as markers for high-risk benign breast disease and incident cancers. Radiology 1980;134:293–5.
18. Ernster VL, Sacks ST, Peterson CA, Schweitzer RJ. Mammographic parenchymal patterns and risk factors for breast cancer. Radiology 1980;134:617–20.
19. Carlile T, Kopecky KJ, Thompson DJ, et al. Breast cancer prediction and the Wolfe classification of mammograms. JAMA 1985;254:1050–3.
20. Saftlas AF, Wolfe JN, Hoover RN, et al. Mammographic parenchymal patterns as indicators of breast cancer risk. Am J Epidemiol 1989;129:518–26.

21. Saftlas AF, Hoover RN, Brinton LA, et al. Mammographic densities and risk of breast cancer. Cancer 1991;67:2833–8.
22. Byrne C, Schairer C, Brinton LA, et al. Effects of mammographic density and benign breast disease on breast cancer risk (United States). Cancer Causes Control 2001;12:103–10.
23. Krook PM, Carlile T, Bush W, Hall MH. Mammographic parenchymal patterns as a risk indicator for prevalent and incident cancer. Cancer 1978;41: 1093–7.
24. Threatt B, Norbeck JM, Ullman NS, Kummer R, Roselle P. Association between mammographic parenchymal pattern classification and incidence of breast cancer. Cancer 1980;45:2550–6.
25. Whitehead J, Carlile T, Kopecky KJ, et al. Wolfe mammographic parenchymal patterns. A study of the masking hypothesis of Egan and Mosteller. Cancer 1985;56:1280–6.
26. Herrinton LJ, Saftlas AF, Stanford JL, Brinton LA, Wolfe JN. Do alcohol intake and mammographic densities interact in regard to the risk of breast cancer? Cancer 1993;71:3029–35.
27. 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.
28. Boyd NF, O'Sullivan B, Campbell JE, et al. Mammographic signs as risk factors for breast cancer. Br J Cancer 1982;45:185–93.
29. Maskarinec G, Meng L. A case-control study of mammographic densities in Hawaii. Breast Cancer Res Treat 2000;63:153–61.
30. 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.
31. Wolfe JN, Saftlas AF, Salane M. Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study. AJR Am J Roentgenol 1987;148:1087–92.
32. Verbeek AL, Hendriks JH, Peeters PH, Sturmans F. Mammographic breast pattern and the risk of breast cancer. Lancet 1984;1:591–3.
33. de Stavola BL, Gravelle IH, Wang DY, et al. Relationship of mammographic parenchymal patterns with breast cancer risk factors and risk of breast cancer in a prospective study. Int J Epidemiol 1990;19:247–54.
34. Ciatto S, Zappa M. A prospective study of the value of mammographic patterns as indicators of breast cancer risk in a screening experience. Eur J Radiol 1993;17:122–5.
35. Byng JW, Yaffe MJ, Lockwood GA, et al. Automated analysis of mammographic densities and breast carcinoma risk. Cancer 1997;80:66–74.
36. Yaffe M, Boyd N, Byng J, et al. Breast cancer risk and measured mammographic density. Eur J Cancer Prev 1998;7:S47–55.
37. Kato I, Beinart C, Bleich A, et al. A nested case-control study of mammographic patterns, breast volume, and breast cancer (New York City, NY, United States). Cancer Causes Control 1995;6:431–8.
38. Thurfjell E, Hsieh CC, Lipworth L, et al. Breast size and mammographic pattern in relation to breast cancer risk. Eur J Cancer Prev 1996;5: 37–41.
39. Salminen TM, Saarenmaa IE, Heikkila MM, Hakama M. Risk of breast cancer and changes in mammographic parenchymal patterns over time. Acta Oncol 1998;37:547–51.
40. van Gils CH, Hendriks JH, Otten JD, Holland R, Verbeek AL. Parity and mammographic breast density in relation to breast cancer risk: indication of interaction. Eur J Cancer Prev 2000;9:105–11.
41. Ziv E, Tice J, Smith-Bindman R, et al. Mammographic density and estrogen receptor status of breast cancer. Cancer Epidemiol Biomarkers Prev 2004;13: 2090–5.
42. 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.
43. 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.
44. Ciatto S, Visioli C, Paci E, Zappa M. Breast density as a determinant of interval cancer at mammographic screening. Br J Cancer 2004;90:393–6.
45. Hainline S, Myers L, McLelland R, et al. Mammographic patterns and risk of breast cancer. Am J Roentgenol 1978;130:1157–8.
46. Weich J, Adler O. Breast pattern as a risk for development of cancer. Isr J Med Sci 1981;17:854–88.
47. Brisson J, Merletti F, Sadowsky NL, et al. Mammographic features of the breast and breast cancer risk. Am J Epidemiol 1982;115:428–37.
48. Janzon L, Andersson I, Petersson F. Mammographic patterns as indicators of risk of breast cancer. Radiology 1982;143:417–9.
49. Tabar L, Dean PB. Mammographic parenchymal patterns. Risk indicator for breast cancer? JAMA 1982;247:185–9.
50. Chaudary M, Gravelle I, Bulstrode J, Wang D, Millis I. Breast parenchymal patterns in women with bilateral primary breast cancer. Br J Radiol 1983;56: 703–6.
51. Brisson J, Morrison AS, Kopans DB, et al. Height and weight, mammographic features of breast tissue, and breast cancer risk. Am J Epidemiol 1984;119:371–81.
52. Horwitz R, Lamas A, Peck D. Mammographic parenchymal patterns and risk of breast cancer in postmenopausal women. Am J Med 1984;77: 621–4.
53. 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.
54. Jakes RW, Duffy SW, Ng FC, Gao F, Ng EH. Mammographic parenchymal patterns and risk of breast cancer at and after a prevalence screen in Singaporean women. Int J Epidemiol 2000;29:11–9.
55. Lam PB, Vacek PM, Geller BM, Muss HB. The association of increased weight, body mass index, and tissue density with the risk of breast carcinoma in Vermont. Cancer 2000;89:369–75.
56. 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.
57. Nagao Y, Kawaguchi Y, Sugiyama Y, Saji S, Kashiki Y. Relationship between mammographic density and the risk of breast cancer in Japanese women: a case-control study. Breast Cancer 2003;10:228–33.
58. Kerlikowske K, Shepherd J, Creasman J, et al. Are breast density and bone mineral density independent risk factors for breast cancer? J Natl Cancer Inst 2005;97:368–74.
59. Nagata C, Matsubara T, Fujita H, et al. Mammographic density and the risk of breast cancer in Japanese women. Br J Cancer 2005;92:2102–6.
60. Egan RL, McSweeney MB. Mammographic parenchymal patterns and risk of breast cancer. Radiology 1979;133:65–70.
61. Kojima O, Majima T, Uehara Y, et al. Radiographic parenchymal patterns in Japanese females as a risk factor for breast carcinoma. World J Surg 1984;8: 414–8.
62. Wolfe J. Breast patterns as an index of risk for developing breast cancer. Am J Roentgenol 1976;126:1130–9.
63. Peyster R, Kalisher L, Cole P. Mammographic parenchymal patterns and the prevalence of breast cancer. Radiology 1977;125:387–91.
64. Boyd NF, O'Sullivan B, Campbell JE, et al. Bias and the association of mammographic parenchymal patterns with breast cancer. Br J Cancer 1982; 45:179–84.
65. Brebner D, Epstein E, Lange M. Xerographic parenchymal patterns and breast cancer. S Afr Med J 1978;54:853–6.
66. Boyd NF, O'Sullivan B, Campbell JE, et al. Mammographic patterns and bias in breast cancer detection. Radiology 1982;143:671–4.
67. D'Orsi CJ, Basset LW, Feig SA, et al. Breast imaging reporting and data systems (BIRADS), 3rd ed. Reston, VA: American College of Radiology; 1998.
68. Boyd NF, Rommens JM, Vogt K, et al. Mammographic breast density as an intermediate phenotype for breast cancer. Lancet Oncol 2005;6:798–808.
69. Boyd NF, Dite GS, Stone J, et al. Heritability of mammographic density, a risk factor for breast cancer. N Engl J Med 2002;347:886–94.
70. Berube S, Diorio C, Verhoek-Oftedahl W, Brisson J. Vitamin D, calcium, and mammographic breast densities. Cancer Epidemiol Biomarkers Prev 2004; 13:1466–72.
71. Boyd NF, Greenberg C, Lockwood G, et al. Effects at two years of a low-fat, high-carbohydrate diet on radiologic features of the breast: results from a randomized trial. Canadian Diet and Breast Cancer Prevention Study Group. J Natl Cancer Inst 1997;89:488–96.
72. 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.
73. Pike MC, Krailo MD, Henderson BE, Casagrande JT, Hoel DG. ''Hormonal'' risk factors, ''breast tissue age,'' and the age-incidence of breast cancer. Nature 1983;303:767–70.