Am J Cancer Res 2023;13(12):6280-6289
www.ajcr.us /ISSN:2156-6976/ajcr0153321

Original Article

Associations of stem cell markers in benign breast tissue with subsequent breast cancer risk

Lusine Yaghjyan[1], Yujing J Heng[2], Gabrielle M Baker[2], Divya Murthy[3], Matt B Mahoney[3], Bernard Rosner[3], Rulla M Tamimi[4]

[1]Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, Gainesville, FL, USA; [2]Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA; [3]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA; [4]Department of Population Health Sciences, Weill Cornell Medicine, New York, NY, USA

Received September 6, 2023; Accepted November 6, 2023; Epub December 15, 2023; Published December 30, 2023

Abstract: We examined associations of stem cell markers CD44, CD24, and ALDH1A1 in benign breast biopsy samples with subsequent breast cancer (BCa) risk and explored if these associations were mediated by mammographic breast density (MBD). We included 101 BCa cases/375 controls, all with previous biopsy-confirmed benign breast disease (BBD) within the Nurses' Health Study (NHS) and NHSII. The data on BCa risk factors were obtained from biennial questionnaires. MBD was assessed with computer-assisted techniques. Immunohistochemistry (IHC) was done on BBD tissue microarrays. For each core, the IHC expression was assessed using a semi-automated method, and expressed as % of cells that stained positive for a specific marker out of the total cell count. Logistic regression was used to examine the associations of each marker's expression of each (in epithelium and stroma) with BCa risk, adjusted for risk factors. Stromal CD44 expression was inversely associated with BCa risk (OR for ≥10% vs. <10%=0.58, 95% CI 0.34, 1.00). Combined stromal + epithelial CD24 expression was inversely associated with BCa risk (>50% vs. 0-10% OR=0.17, 95% CI 0.04-0.81, p-trend =0.03). Stromal CD24 and ALDH1A1 as well as epithelial expression of any of the three markers were not associated with BCa risk. In a smaller subset of women with available MBD, these observed associations did not appear to be mediated by MBD. Our findings suggest inverse associations of CD44 in stroma and combined stromal + epithelial CD24 with BCa risk. Future studies are warranted to confirm our findings and to examine these associations by BBD subtype.

Keywords: Benign breast disease, breast cancer risk, stem cell markers

Introduction

Breast tissue undergoes significant structural changes throughout the woman's life [1]. The tissue architecture is maintained by a population of stem cells with self-renewal capacity, which are essential for tissue repair and remodeling [2]. However, potentially limitless self-renewal capacity and high susceptibility to various endogenous and exogenous mutagenic insults increase the chances of their tumorigenic transformation [1, 3]. A stem cell hypothesis of breast carcinogenesis suggests that the breast cancer development might be directly related to the size of the stem cell pool and its mitotic activity [4]. Further, in the mammary gland, stem cells are the only cell subpopulation that can accumulate all the oncogenic alterations [1].

Well-characterized stem cell markers CD44 and CD24 have been linked to younger age at diagnosis, higher odds of unfavorable tumor characteristics, including triple-negative state, and distant metastasis [5-8]. Another stem cell marker, aldehyde dehydrogenase family 1 member A1 (ALDH1A1), is correlated with poor prognosis and chemotherapy resistant breast cancer [7, 9-13]. However, due to the very limited availability of pre-diagnostic healthy breast tissue, the associations of these markers with breast cancer risk have not been explored and the data on the expression of stem cell markers in the breast tissue of cancer-free women is

very limited. In our earlier study, using women who previously volunteered for a unique biopsy study of normal breast tissue at the Mayo Clinic (2006-2008), we found positive associations of CD44 expression and suggestive associations of CD24 and ALDH1A1 with mammographic breast density (MBD), a well-established, strong breast cancer risk factor reflective of relative extent of fibroglandular and adipose tissue in the breast [14]. Whether the expression of these markers in benign breast biopsy samples could predict subsequent breast cancer risk is unknown.

To fill this gap, we examined the associations of stem cell markers CD44, CD24, and ALDH1A1 in benign breast biopsy samples with subsequent breast cancer risk using prospective data from the Nurses' Health Study (NHS) and Nurses' Health Study II (NHSII).

Materials and methods

Study population

Our analysis included women with biopsy-confirmed benign breast disease (BBD) in the Nurses' Health Study (NHS) and Nurses' Health Study II (NHSII) cohorts who were previously included in a nested case-control study of breast cancer [15, 16]. These prospective cohorts followed registered nurses in the United States who were 30-55 (NHS) or 25-42 years old (NHSII) at enrollment. After administration of the initial questionnaire, the information on breast cancer risk factors (body mass index [BMI], reproductive history, postmenopausal hormone [PMH] use, and alcohol use) and any diagnoses of cancer or other diseases (including BBD) was updated through biennial questionnaires which were then confirmed via medical record review [17]. Details of this nested case-control study and the BBD assessment have been previously described [15, 16].

Early NHS questionnaires (1976, 1978, and 1980) asked whether the participant had ever been diagnosed with 'fibrocystic disease' or 'other BBD' and whether she had been hospitalized in relation to this diagnosis. Beginning in 1982, the NHS questionnaires specifically asked about a history of biopsy-confirmed BBD (fibrocystic disease or other BBD). The initial 1989 NHS II questionnaire and all subsequent biennial questionnaires also asked participants to report any diagnosis of BBD and to indicate whether it was confirmed by biopsy or aspiration.

Cases were women with biopsy-confirmed BBD who reported a diagnosis of breast cancer during 1976-1998 for the NHS and 1989-1999 for the NHSII following their BBD diagnosis. Using incidence density sampling, four women with biopsy-confirmed BBD who were free of breast cancer at the time of the matching case's diagnosis (controls) were matched to the respective case on year of birth and year of benign breast biopsy [18]. We obtained BBD pathology records and archived biopsy specimens for all cases and controls from their hospital pathology departments; our ability to obtain biopsy blocks did not significantly differ by case and control status. Women were excluded if they had evidence of in situ or invasive carcinoma or unknown lesion type at the time of benign breast biopsy (n=34). All cases and controls from this nested case-control were cancer-free at the time of BBD diagnosis, with an average time of 9 years between biopsy and breast cancer diagnosis date. In the current analysis, we included 101 cases and 375 controls who had complete data on breast cancer factors and staining results for stem cell markers.

The study protocol was approved by the institutional review boards (IRB) of the Brigham and Women's Hospital and Harvard T.H. Chan School of Public Health, and those of participating registries as required, and University of Florida IRB. Consent was obtained or implied by return of questionnaires.

Benign breast biopsy confirmation and BBD subtypes

Hematoxylin and eosin (H&E) breast tissue slides were retrieved for biopsy-confirmed BBD patients who gave permission to review their biopsy records. The slides were previously independently reviewed by one of three pathologists in a blinded fashion, i.e. the evaluating pathologists were blinded to type of BBD noted on the original diagnosis [19, 20]. Any slide identified as having either questionable atypia or atypia was jointly reviewed by two pathologists [19, 20]. For each set of slides, a detailed worksheet was completed and the benign breast biopsy was classified according to the categories of Page et al. [21] as non-prolifera-

tive, proliferative without atypia, or atypical hyperplasia (ductal or lobular hyperplasia) [15].

Tissue microarray (TMA) construction of BBD samples

After centralized review of H&E stained slides, we retrieved archived FFPE benign breast biopsy blocks for participants. H&E sections of the corresponding FFPE tissue blocks were re-reviewed by a single pathologist to identify areas of benign proliferative lesions and normal terminal duct-lobular units (TDLUs), and to identify the areas from which the cores for the TMAs would be taken. Normal TDLUs were regions of histologically normal tissue that may or may not be adjacent to benign lesions (e.g., atypical ductal hyperplasia, usual ductal hyperplasia) [22]. TMAs were constructed at the Dana Farber/Harvard Cancer Center (DF/HCC) Tissue Microarray Core Facility by obtaining 0.6-mm cores from benign lesions and TDLUs. For each woman, up to 3 cores of normal TDLU were included in the TMA blocks. We previously evaluated our TMA construction methods and confirmed a high success rate (76%) of capturing normal TDLUs in these TMA blocks [23].

Immunohistochemistry (IHC) for stem cell markers

The expression of the stem cell markers was evaluated by automated IHC technique that allows the quantification of markers' expression levels and localization of the target signal to specific cells/structures. For each of the three markers one 5-μm paraffin section was cut from a single TMA block and then stained at the University of Florida Pathology Core Lab on DAKO AutostainerPlus according to the previously standardized protocol with commercial antibodies (CD44 [DAKO] 1:25 dilution; CD24 [Invitrogen] 1:200 dilution and ALDH1A1 [Abcam] 1:300 dilution). Details of this protocol have been described previously [24-26]. Briefly, slides were de-paraffinized with xylene and re-hydrated through decreasing concentrations of ethanol to water, including an intermediate step to quench endogenous peroxidase activity (3% hydrogen peroxide in methanol) and transferred to 1X TBS-T (Tris-buffered saline-Tween). For heat-induced antigen retrieval, sections were heated in a steamer while submerged in Citra (Biogenex, Fremont, CA) or Trilogy (Cell Marque, Rocklin, CA) for 30 minutes. Next, slides were 1) rinsed in 1XTBS-T and incubated with a universal protein blocker Sniper (Biocare Medical, Walnut Creek, CA) for 10 (for CD44 and ALDH1A1) or 15 minutes (for CD24); 2) rinsed in 1XTBS-T and co-incubated in primary antibody ALDH1A1, CD24, or CD44 for 1 hour; and 3) rinsed in 1XTBS-T followed by application of conjugated secondary antibody (Mach 2 goat anti-rabbit horse [or mouse] radish peroxidase-conjugated, Biocare Medical, Walnut Creek, CA) for 30 minutes. Detection of antibodies was achieved by incubating slides in 3'3' diaminobenzidine (Vector Laboratories Inc., Burlingame, CA) for 4 minutes. Slides were counterstained with hematoxylin (Biocare Medical, Walnut Creek, CA) 1:10 for 3 minutes and mounted with Cytoseal XYL (Richard-Allen Scientific, Kalamazoo, MI). The laboratory implemented standard quality control procedures.

Image analysis

Immunoreactivity was quantified using a semi-automated image analysis system, Definiens Tissue Studio® software (Munich, Germany) which quantifies tissue marker expression within the context of tissue architecture. In our recent study, we demonstrated that IHC staining quantification is highly comparable across various software applications for IHC analysis (Definiens, InForm®, [Akoya Biosciences, Marlborough, MA] and QuPath) [27]. For each core, the extent of each marker expression was assessed on a continuous scale as percent of cells that stain positively (across all intensities) for a specific marker out of the total cell count, separately for epithelium and stroma. Briefly, TMA slides were digitized at 20× into whole slide images using the Pannoramic Scan 150P (3DHistech, Budapest, Hungary). For each marker, the images were imported into Definiens and a representative TMA randomly selected as the training TMA [27, 28]. On the training TMA, the operator selected 12 training cores that were assessed as >0-<1 (n=3), 1-10 (n=3), >10-50 (n=3), and >50% (n=3) by the pathologist to optimize a Definiens' algorithm for automated IHC assessment. Definiens only allows a maximum of 12 cores for algorithm training. The minimum positive IHC staining threshold in Definiens was set using the pathologist's manual reads as reference. The optimized Definiens algorithm segmented each tis-

sue core into epithelium, stroma, and fat, detected the number of cells, and quantified the IHC expression.

The current analysis was specifically focused on the expression of stem cell markers in normal TDLU cores for the following reasons: (1) we specifically targeted normal TDLUs in construction of these TMAs within NHS/NHSII and thus the number of women with benign lesion cores was smaller and would not allow to draw meaningful conclusions; (2) in our earlier reliability study, we observed higher heterogeneity within benign lesion cores as they were represented by various lesion types [24]; and (3) we were interested in the underlying changes in the breast tissue happening early in the process of breast carcinogenesis and thus normal TDLUs were more relevant to address our research questions.

Staining results for stroma were available for 96 cases/338 controls, 98 cases/347 controls, and 91 cases/328 controls for CD44, CD24, and ALDH1A1, respectively; the staining results for epithelium were available for 93 cases/314 controls, 95 cases/325 controls, and 92 cases/311 controls, for CD44, CD24, and ALDH1A1, respectively.

Covariate information

Information on breast cancer risk factors was obtained from the biennial questionnaires closest to the reference date (date of diagnosis for cases and their matched controls). Women were considered to be postmenopausal if they reported: 1) no menstrual periods within the 12 months before biopsy with natural menopause; 2) bilateral oophorectomy; or 3) hysterectomy with one or both ovaries retained, and were 54 years or older for ever smokers or 56 years or older for never smokers [29, 30]. In additional analysis, we also considered percent mammographic breast density which was measured with computer-assisted determination (the Cumulus software, University of Toronto, Toronto, Canada) [17, 31] and was available for a smaller subset of women. Percent breast density was measured as percentage of the total area occupied by epithelial/stromal tissue (absolute dense area) divided by the total breast area. Because breast densities of the right and left breast for an individual woman are strongly correlated [31], the average density of both breasts was used in this analysis.

Statistical analysis

We used logistic regression to examine the associations of expression of each of the markers, adjusted for the following covariates: age (continuous), BMI (continuous), a family history of breast cancer (yes/no), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/no hormones, postmenopausal/past hormones, postmenopausal/current hormones, postmenopausal/unknown hormone use status, unknown menopausal status), age at menarche (<12, 12, 13, >13, unknown), combined parity/age at first birth (parous with first birth before age 25, parous with first birth at or after age 25, nulliparous, unknown), alcohol use (none, >0-<5, ≥5 g/day), and NHS cohort. We modeled marker expression (weighted average across available cores for a woman) as continuous (per 10% increase in expression), dichotomous using 10% cutoffs based on the results of our prior reliability study and distribution in our sample [24], as well as categorical (0-10%, >10-50%, and >50%). In additional models, we also adjusted the estimates for BBD subtype. Finally, as we previously reported associations of stem cell markers with mammographic breast density in NHS/NHSII (manuscript under review), in a smaller subset of women with available mammographic breast density, we additionally adjusted the models for percent breast density (continuous) and assessed whether and to what extent the associations between each marker and breast cancer risk may be mediated by percent density. We implemented the method for mediation analysis outlined by Lin et al. using the SAS macro developed by Spiegelman and colleagues at the Harvard T. H. Chan School of Public Health [32]. Additional information on the method and SAS macro can be found at: http://www.hsph.harvard.edu/donna-spiegelman/software/mediate/. Using this method, we estimated the β coefficients and 95% confidence intervals (CIs) for (a) the association between the marker and breast cancer risk not adjusted for PMD and (b) the association between the marker and breast cancer risk adjusted for percent density. We then estimated the percent of the total association between the marker and breast cancer risk

Table 1. Age-adjusted characteristics of study participants, by case-control status

Characteristic	Controls n=375	Cases n=101	P for difference
Mean (SD)			
Age (years)[a]	52.47 (8.62)	53.92 (8.68)	NA
Age at menarche (years)	12.58 (1.37)	12.27 (1.43)	0.35
Body Mass Index (kg/m^2)	24.77 (4.66)	23.88 (3.73)	0.91
Age at menopause (years)	49.06 (5.19)	50.15 (2.31)	0.61
Alcohol, g/day	5.166 (8.78)	3.438 (4.22)	0.83
CD44 normal TDLU Epithelium %	41.24 (36.58)	38.04 (34.12)	0.13
CD44 normal TDLU Stroma %	19.09 (27.92)	17.09 (26.4)	0.24
CD24 normal TDLU Epithelium %	29.32 (23.13)	28.44 (21.54)	0.28
CD24 normal TDLU Stroma %	9.001 (14.423)	7.122 (9.831)	0.88
ALDH1A1 normal TDLU Epithelium %	27.21 (19.18)	27.52 (20.84)	0.03
ALDH1A1 normal TDLU Stroma %	12 (13.8)	12.32 (13.14)	0.29
Percentages			
Parity/age at first birth			0.57
Nulliparous	9	8	
Parous, age <25 years	50	43	
Parous, age ≥25 years	40	48	
Family history of breast cancer	11	24	0.32
Benign breast disease			0.18
Non-proliferative	31	24	
Proliferative without atypia	56	51	
Proliferative with atypia	13	25	
Never smoked	49	36	0.50
Past smoker	5	6	
Current smoker	13	17	
Premenopausal	40	39	0.82
Postmenopausal/never used MHT	25	27	
Postmenopausal/past MHT	14	14	
Postmenopausal/current MHT	19	16	

[a]Value is not age adjusted.

that was mediated by percent density using the following equation: PctMed = (1 - (estimate for adjusted/estimate for unadjusted)) * 100, and the corresponding *p*-value for mediation. All the analyses were performed using SAS software (version 9.4, SAS Institute, Cary, NC). All tests of statistical significance were 2-sided.

Results

In this study of 476 women (101 cases and 375 controls), 141 (29.6%) had non-proliferative disease, 257 (54.0%) had proliferative disease without atypia, and 78 (16.4%) had atypical hyperplasia, consistent with previously reported distributions of these BBD subtypes [20]. The average age at diagnosis (reference date) was 53 years (range 29-73 years). Majority of the women were postmenopausal at the reference date (59.6%). Age-adjusted characteristics of women in the study by case-control status are presented in **Table 1**. Compared to controls, cases on average consumed less alcohol (3.4 vs. 5.2 g/day), had lower BMI (23.9 kg/m^2 vs. 24.8 kg/m^2), and were more likely to have a family history of breast cancer (24% vs. 11%) and to have proliferative BBD with atypia (25% vs. 13%). Distribution of stem cell markers' expression by case-control status and BBD subtype are presented in Supplementary Table 1. Among both cases and controls, women with proliferative BBD with atypia had greater stro-

Table 2. Associations of stem cell markers' expression in normal terminal duct-lobular units (weighted average across cores) with breast cancer risk (odds ratios and 95% confidence intervals)[a]

Stem cell marker expression	CD44		CD24		ALDH1A1	
	N cases/ controls	OR (95% CI)	N cases/ controls	OR (95% CI)	N cases/ controls	OR (95% CI)
Stroma						
Continuous, per 10% increase	94/324	0.95 (0.87-1.05)	95/333	0.87 (0.71-1.06)	90/315	0.92 (0.76-1.11)
<10%	67/197	Ref	71/248	Ref	59/191	Ref
≥10%	27/127	0.58 (0.34-1.00)	24/85	0.85 (0.48-1.49)	31/124	0.81 (0.49-1.36)
0-10%	67/197	Ref	71/249	Ref	59/191	Ref
>10-50%	13/76	0.47 (0.23-0.94)	24/67	1.12 (0.63-1.99)	29/114	0.86 (0.51-1.46)
>50%	14/51	0.75 (0.37-1.51)	0/17	NE	2/10	0.37 (0.06-2.26)
p-trend	94/324	0.28	95/333	0.10	90/315	0.24
Epithelium						
Continuous, per 10% increase	91/300	0.96 (0.90-1.03)	92/312	0.96 (0.86-1.08)	90/299	1.01 (0.89-1.16)
<10%	30/93	Ref	26/63	Ref	26/67	Ref
≥10%	61/207	1.03 (0.61-1.76)	66/249	0.67 (0.38-1.19)	64/132	0.66 (0.37-1.19)
0-10%	30/93	Ref	26/63	Ref	26/67	Ref
>10-50%	35/96	1.42 (0.77-2.59)	48/188	0.66 (0.36-1.20)	52/196	0.63 (0.34-1.14)
>50%	26/111	0.73 (0.39-1.38)	18/61	0.69 (0.32-1.47)	12/36	0.86 (0.36-2.02)
p-trend	91/300	0.17	92/312	0.48	90/299	0.70
Stroma + epithelial						
Continuous, per 10% change	94/325	0.96 (0.88-1.04)	96/335	0.94 (0.81-1.08)	96/326	0.99 (0.83-1.16)
<10%	43/144	Ref	38/129	Ref	29/90	Ref
≥10%	51/181	1.06 (0.64-1.75)	58/206	0.90 (0.55-1.48)	67/236	0.79 (0.46-1.38)
0-10%	44/144	Ref	38/129	Ref	29/91	Ref
>10-50%	28/99	1.17 (0.66-2.09)	56/177	1.03 (0.62-1.70)	64/220	0.82 (0.47-1.42)
>50%	22/82	0.83 (0.44-1.56)	2/29	0.17 (0.04-0.81)	3/15	0.61 (0.15-2.43)
p-trend	94/325	0.51	96/335	0.03	96/326	0.39

Abbreviations: CI, confidence interval; OR, odds ratio; NE, not estimable. [a]Adjusted for age (continuous), BMI (continuous), a family history of breast cancer (Yes/No), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/no hormones, postmenopausal/past hormones, postmenopausal/current hormones, postmenopausal/unknown hormone use status, unknown menopausal status), age at menarche (<12, 12, 13, >13, unknown), combined parity/age at first birth (parous with first birth before age 25, parous with first birth at or after age 25, nulliparous, unknown), alcohol use (none, >0-<5, ≥5 g/day), and NHS cohort.

mal and epithelial expression in normal TDLUs for all three markers as compared to non-proliferative disease and proliferative without atypia. Cases had lower expression of all three markers across all BBD subtypes (Supplementary Table 1).

In multivariate analysis (**Table 2** and Supplementary Figure 1), stromal CD44 expression in normal TDLUs was inversely associated with breast cancer risk (OR for ≥10% vs. <10%=0.58, 95% CI 0.34, 1.00). Stromal expression of CD24 and ALDH1A1 was not associated with breast cancer risk (OR=0.85, 95% CI 0.48, 1.49 and OR=0.81, 95% CI 0.49, 1.36, respectively). We found no association of epithelial expression of any of the three markers with breast cancer risk (p-trend >0.17 for all). Combined stromal + epithelial CD24 expression was inversely associated with breast cancer risk (>50% vs. 0-10% OR=0.17, 95% CI 0.04-0.81, p-trend =0.03) (**Table 2**). These findings remained similar after additional adjustment for BBD subtype (Supplementary Table 2).

In a smaller subset of women with available mammographic breast density (21 cases/160 controls), additional adjustment of the risk estimates for percent density resulted in stronger inverse associations of stromal CD44 expression with breast cancer risk (OR for ≥10% vs. <10%=0.20, 95% CI 0.05, 0.86) (Supplementary Table 3). Mediation analysis revealed that these associations were not mediated by breast density. While other markers' expres-

sion measures were not associated with breast cancer risk after adjustment for percent breast density, we found a suggestive evidence of mediation by percent density for associations of combined stromal/epithelial CD24 expression with breast cancer risk (percent mediated: 30.7%, P=0.21). While suggestive evidence of mediation was also found for stromal CD24 as well as epithelial and combined stromal/epithelial ALDH1A1 expression, the percent mediated was <7.7% for all. There was no evidence of mediation by percent density for associations of other marker expression measures and breast cancer risk (Supplementary Table 3).

Discussion

In this study of 101 cases and 375 controls, we found inverse associations of stromal CD44 expression and combined stromal + epithelial CD24 expression with breast cancer risk. Associations for CD44 became stronger after additional adjustment for percent breast density. Expression of ALDH1A1 was not associated with breast cancer risk.

The data on expression of breast stem cell markers CD44, CD24, and ALDH1A1 in non-cancerous breast tissue is very limited. In previous studies with tumor tissue, these markers have been linked to younger age at diagnosis, unfavorable tumor characteristics (i.e. grade, stage, and triple-negative status), metastatic spread, poor prognosis and chemotherapy resistance [5-13], with positive associations for CD44 and ALDH1A1 and inverse associations for CD24. In our study, in contrast to our hypothesis of positive associations of CD44 and ALDH1A1 and inverse associations of CD24 with breast cancer risk (based on the evidence from studies in tumors), we found an inverse association of stromal CD44 expression in normal TDLUs with breast cancer risk.

Cancers in epithelial organs, including the breast, may result from deregulation of normal stem-cell functions, including self-renewal, ability to differentiate, active telomerase and antiapoptotic pathways, increased membrane transporter activity, and anchorage independence all of which are normally tightly regulated [33]. Disruption in these normal processes could lead to development of pro-tumorigenic phenotypes and, eventually, breast tumors [33]. Some studies speculate that stem cells in the normal breast epithelium might be located in the basal layer and that these cells might be early descendants of breast epithelial stem cells resulting from deregulation of normal processes and expanded in benign proliferative phenotypes [34]. However, we found no association of epithelial expression of any of the three markers with breast cancer risk, with and without adjustment for BBD subtype.

Previous studies suggest that CD44(+)CD24(-/low) and ALDH1(high) expression could be used to characterize two largely non-overlapping populations of breast cancer stem cells which have epithelial-like and mesenchymal-like phenotypes, respectively [35-37]. However, less is known about these cells and application of these markers in non-cancerous breast tissue. In our study, we were unable to assess combination of these markers' expression on a cell-by-cell basis. Additionally, we specifically targeted normal TDLUs to be able to capture changes that may occur early in breast carcinogenesis. As we observed an overall greater expression of these markers in the normal TDLUs of women with proliferative benign breast disease with atypia compared to non-proliferative disease or proliferative disease without atypia, it is possible that the marker expression changes become apparent at later stages and more pronounced in the lesion areas. However, as TMA construction in NHS/NHSII specifically targeted normal TDLUs, we were underpowered to examine associations of markers' expression in lesion cores with breast cancer risk. Importantly, in our earlier reliability study, we also observed higher heterogeneity within benign lesion cores as they were represented by various lesion types (atypical ductal hyperplasia, atypical lobular hyperplasia, apocrine metaplasia, non-apocrine cysts, and usual ductal hyperplasia). Thus, future studies specifically targeting various subtypes of benign lesions for TMA construction are needed to further understand these associations.

To our knowledge, this is the first study to date exploring associations of breast stem cell markers CD44, CD24, and ALDH1A1 with breast cancer risk. The analysis used data from the Nurses' Health Study and Nurses' Health Study II, established cohorts with more than 30 years of follow-up, confirmed benign breast disease status, and comprehensive information

on breast cancer risk factors. Our study has a few limitations. We recognize that biopsy samples come from a specific area of the breast. Our previous work demonstrates that this sampling approach still provides strong evidence for a priori hypotheses and meaningful findings for breast tissue involution [38], identification of markers associated with breast cancer risk [22, 39, 40], and associations with known breast cancer risk factors, suggesting that this limitation has minimal impact on research findings [41]. Next, as we did not use co-localization of the markers during IHC in our study, were unable to assess combination of these markers' expression on a cell-by-cell basis. Finally, we were underpowered to examine associations stratified by the BBD subtype and by menopausal status.

In conclusion, we investigated the associations of the expression of stem cell markers CD44, CD24, and ALDH1A1 with subsequent breast cancer risk in women with benign breast biopsies. Our findings suggest inverse associations of CD44 expression in stroma with breast cancer risk. Future studies are warranted to confirm our findings and to examine these associations by BBD subtype.

Acknowledgements

This work was supported by the National Cancer Institute at the National Institutes of Health (CA240341 to L.Y., CA131332, CA17-5080, P01 CA087969 to R.M.T., UM1 CA186107 to M.S., and U01 CA176726 to W.W.), Avon Foundation for Women, Susan G. Komen for the Cure®, and Breast Cancer Research Foundation. The authors would like to acknowledge the contribution to this study from central cancer registries supported through the Centers for Disease Control and Prevention's National Program of Cancer Registries (NPCR) and/or the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program. Central registries may also be supported by state agencies, universities, and cancer centers. Participating central cancer registries include the following: Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Indiana, Iowa, Kentucky, Louisiana, Massachusetts, Maine, Maryland, Michigan, Mississippi, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Puerto Rico, Rhode Island, Seattle SEER Registry, South Carolina, Tennessee, Texas, Utah, Virginia, West Virginia, Wyoming.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Lusine Yaghjyan, Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, 2004 Mowry Rd., Gainesville, FL 32610, USA. Tel: 352-294-5940; ORCID: 0000-0002-1626-5340; Fax: 352-273-5365; E-mail: lyaghjyan@ufl.edu

References

[1] Cobaleda C, Cruz JJ, Gonzalez-Sarmiento R, Sanchez-Garcia I and Perez-Losada J. The emerging picture of human breast cancer as a stem cell-based disease. Stem Cell Rev 2008; 4: 67-79.

[2] Smalley M and Ashworth A. Stem cells and breast cancer: a field in transit. Nat Rev Cancer 2003; 3: 832-844.

[3] Chang CC. Recent translational research: stem cells as the roots of breast cancer. Breast Cancer Res 2006; 8: 103.

[4] Ginestier C and Wicha MS. Mammary stem cell number as a determinate of breast cancer risk. Breast Cancer Res 2007; 9: 109.

[5] Giatromanolaki A, Sivridis E, Fiska A and Koukourakis MI. The CD44+/CD24- phenotype relates to 'triple-negative' state and unfavorable prognosis in breast cancer patients. Med Oncol 2011; 28: 745-52.

[6] Abraham BK, Fritz P, Kuip HVD, Buck M, Szabo Z, Athelogou M and Brauch H. Evaluation of CD44+/CD24-/low cells in breast cancer and relevance for distant metastasis. AACR Meeting Abstracts 2005; 65: 481.

[7] Neumeister V, Agarwal S, Bordeaux J, Camp RL and Rimm DL. In situ identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis. Am J Pathol 2010; 176: 2131-2138.

[8] Lee JH, Kim SH, Lee ES and Kim YS. CD24 overexpression in cancer development and progression: a meta-analysis. Oncol Rep 2009; 22: 1149-1156.

[9] Eden JA. Human breast cancer stem cells and sex hormones--a narrative review. Menopause 2010; 17: 801-810.

[10] Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, Jacquemier J, Viens P, Kleer CG, Liu S, Schott A, Hayes D, Birnbaum D, Wicha MS and Dontu G. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell 2007; 1: 555-567.
[11] Liu Y, Lv DL, Duan JJ, Xu SL, Zhang JF, Yang XJ, Zhang X, Cui YH, Bian XW and Yu SC. ALDH1A1 expression correlates with clinicopathologic features and poor prognosis of breast cancer patients: a systematic review and meta-analysis. BMC Cancer 2014; 14: 444.
[12] Khoury T, Ademuyiwa FO, Chandrasekhar R, Jabbour M, DeLeo A, Ferrone S, Wang Y and Wang X. Aldehyde dehydrogenase 1A1 expression in breast cancer is associated with stage, triple negativity, and outcome to neoadjuvant chemotherapy. Mod Pathol 2012; 25: 388-397.
[13] Yao J, Jin Q, Wang XD, Zhu HJ and Ni QC. Aldehyde dehydrogenase 1 expression is correlated with poor prognosis in breast cancer. Medicine (Baltimore) 2017; 96: e7171.
[14] Yaghjyan L, Stoll E, Ghosh K, Scott CG, Jensen MR, Brandt KR, Visscher D and Vachon CM. Tissue-based associations of mammographic breast density with breast stem cell markers. Breast Cancer Res 2017; 19: 100.
[15] Tamimi RM, Colditz GA, Wang Y, Collins LC, Hu R, Rosner B, Irie HY, Connolly JL and Schnitt SJ. Expression of IGF1R in normal breast tissue and subsequent risk of breast cancer. Breast Cancer Res Treat 2011; 128: 243-250.
[16] Collins LC, Baer HJ, Tamimi RM, Connolly JL, Colditz GA and Schnitt SJ. The influence of family history on breast cancer risk in women with biopsy-confirmed benign breast disease: results from the Nurses' Health Study. Cancer 2006; 107: 1240-1247.
[17] Tamimi RM, Byrne C, Colditz GA and Hankinson SE. Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women. J Natl Cancer Inst 2007; 99: 1178-1187.
[18] Tamimi RM, Rosner B and Colditz GA. Evaluation of a breast cancer risk prediction model expanded to include category of prior benign breast disease lesion. Cancer 2010; 116: 4944-4953.
[19] Jacobs TW, Byrne C, Colditz G, Connolly JL and Schnitt SJ. Pathologic features of breast cancers in women with previous benign breast disease. Am J Clin Pathol 2001; 115: 362-369.
[20] Tamimi RM, Byrne C, Baer HJ, Rosner B, Schnitt SJ, Connolly JL and Colditz GA. Benign breast disease, recent alcohol consumption, and risk of breast cancer: a nested case-control study. Breast Cancer Res 2005; 7: R555-62.
[21] Page DL, Dupont WD, Rogers LW and Rados MS. Atypical hyperplastic lesions of the female breast. A long-term follow-up study. Cancer 1985; 55: 2698-2708.
[22] Tamimi RM, Colditz GA, Wang Y, Collins LC, Hu R, Rosner B, Irie HY, Connolly JL and Schnitt SJ. Expression of IGF1R in normal breast tissue and subsequent risk of breast cancer. Breast Cancer Res Treat 2011; 128: 243-250.
[23] Collins LC, Wang Y, Connolly JL, Baer HJ, Hu R, Schnitt SJ, Colditz GA and Tamimi RM. Potential role of tissue microarrays for the study of biomarker expression in benign breast disease and normal breast tissue. Appl Immunohistochem Mol Morphol 2009; 17: 438-441.
[24] Yaghjyan L, Heng YJ, Baker GM, Bret-Mounet V, Murthy D, Mahoney MB, Mu Y, Rosner B and Tamimi RM. Reliability of CD44, CD24, and ALDH1A1 immunohistochemical staining: pathologist assessment compared to quantitative image analysis. Front Med (Lausanne) 2022; 9: 1040061.
[25] Yaghjyan L, Stoll E, Ghosh K, Scott CG, Jensen MR, Brandt KR, Visscher D and Vachon CM. Tissue-based associations of mammographic breast density with breast stem cell markers. Breast Cancer Res 2017; 19: 100.
[26] Yaghjyan L, Esnakula AK, Scott CG, Wijayabahu AT, Jensen MR and Vachon CM. Associations of mammographic breast density with breast stem cell marker-defined breast cancer subtypes. Cancer Causes Control 2019; 30: 1103-1111.
[27] Baker GM, Bret-Mounet VC, Wang T, Veta M, Zheng H, Collins LC, Eliassen AH, Tamimi RM and Heng YJ. Immunohistochemistry scoring of breast tumor tissue microarrays: a comparison study across three software applications. J Pathol Inform 2022; 13: 100118.
[28] Roberts MR, Baker GM, Heng YJ, Pyle ME, Astone K, Rosner BA, Collins LC, Eliassen AH and Tamimi RM. Reliability of a computational platform as a surrogate for manually interpreted immunohistochemical markers in breast tumor tissue microarrays. Cancer Epidemiol 2021; 74: 101999.
[29] Willett W, Stampfer MJ, Bain C, Lipnick R, Speizer FE, Rosner B, Cramer D and Hennekens CH. Cigarette smoking, relative weight, and menopause. Am J Epidemiol 1983; 117: 651-658.
[30] Stampfer MJ, Willett WC, Colditz GA, Rosner B, Speizer FE and Hennekens CH. A prospective study of postmenopausal estrogen therapy and coronary heart disease. N Engl J Med 1985; 313: 1044-1049.

[31] Byng JW, Boyd NF, Little L, Lockwood G, Fishell E, Jong RA and Yaffe MJ. Symmetry of projection in the quantitative analysis of mammographic images. Eur J Cancer Prev 1996; 5: 319-327.

[32] Lin DY, Fleming TR and De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. Stat Med 1997; 16: 1515-1527.

[33] Dontu G, Al-Hajj M, Abdallah WM, Clarke MF and Wicha MS. Stem cells in normal breast development and breast cancer. Cell Prolif 2003; 36 Suppl 1: 59-72.

[34] Petersen OW and Polyak K. Stem cells in the human breast. Cold Spring Harb Perspect Biol 2010; 2: a003160.

[35] Colacino JA, Azizi E, Brooks MD, Harouaka R, Fouladdel S, McDermott SP, Lee M, Hill D, Madden J, Boerner J, Cote ML, Sartor MA, Rozek LS and Wicha MS. Heterogeneity of human breast stem and progenitor cells as revealed by transcriptional profiling. Stem Cell Reports 2018; 10: 1596-1609.

[36] Ricardo S, Vieira AF, Gerhard R, Leitão D, Pinto R, Cameselle-Teijeiro JF, Milanezi F, Schmitt F and Paredes J. Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. J Clin Pathol 2011; 64: 937-946.

[37] Escudero Mendez L, Srinivasan M, Hamouda RK, Ambedkar B, Arzoun H, Sahib I, Fondeur J and Mohammed L. Evaluation of CD44+/CD24- and aldehyde dehydrogenase enzyme markers in cancer stem cells as prognostic indicators for triple-negative breast cancer. Cureus 2022; 14: e28056.

[38] Rice MS, Tamimi RM, Connolly JL, Collins LC, Shen D, Pollak MN, Rosner B, Hankinson SE and Tworoger SS. Insulin-like growth factor-1, insulin-like growth factor binding protein-3 and lobule type in the Nurses' Health Study II. Breast Cancer Res 2012; 14: R44.

[39] Huh SJ, Oh H, Peterson MA, Almendro V, Hu R, Bowden M, Lis RL, Cotter MB, Loda M, Barry WT, Polyak K and Tamimi RM. The proliferative activity of mammary epithelial cells in normal tissue predicts breast cancer risk in premenopausal women. Cancer Res 2016; 76: 1926-1934.

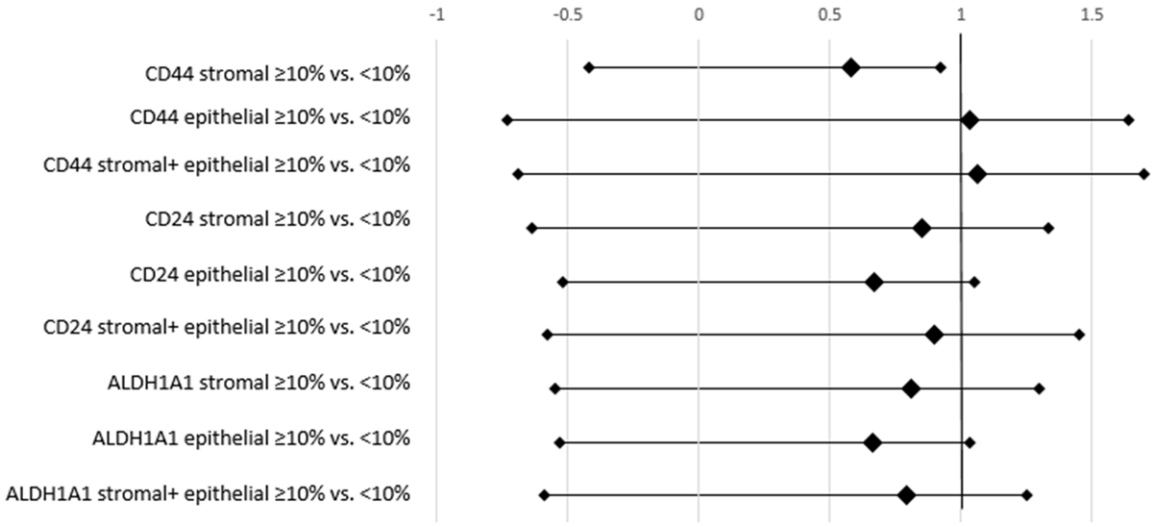
[40] Oh H, Eliassen AH, Wang M, Smith-Warner SA, Beck AH, Schnitt SJ, Collins LC, Connolly JL, Montaser-Kouhsari L, Polyak K and Tamimi RM. Expression of estrogen receptor, progesterone receptor, and Ki67 in normal breast tissue in relation to subsequent risk of breast cancer. NPJ Breast Cancer 2016; 2: 16032.

[41] Oh H, Eliassen AH, Beck AH, Rosner B, Schnitt SJ, Collins LC, Connolly JL, Montaser-Kouhsari L, Willett WC and Tamimi RM. Breast cancer risk factors in relation to estrogen receptor, progesterone receptor, insulin-like growth factor-1 receptor, and Ki67 expression in normal breast tissue. NPJ Breast Cancer 2017; 3: 39.

Supplementary Table 1. Distribution of stem cell markers in normal terminal duct-lobular units by case-control status and benign breast disease subtype (mean [SD] and range)

Stem cell marker expression	BBD subtype			Cases			Controls		
	Non-proliferative n=141	Proliferative without atypia n=257	Proliferative with atypia n=78	Non-proliferative n=24	Proliferative without atypia n=48	Proliferative with atypia n=29	Non-proliferative n=117	Proliferative without atypia n=209	Proliferative with atypia n=29
CD44 stroma	18.6 (28.5) 0-96.2	17.4 (27.4) 0-94.9	21.2 (27.7) 0-93.8	14.3 (26.9) 0-80.8	16.5 (28.1) 0-90.5	17.6 (28.2) 0-93.6	19.5 (28.9) 0-96.2	17.6 (37.3) 0-94.9	23.5 (27.5) 0-94.8
CD44 epithelium	27.4 (37.4) 0-100	38.4 (35.9) 0-100	49.0 (33.1) 0.9-100	30.5 (36.1) 0-98.9	35.6 (33.6) 0-98.8	42.8 (32.5) 0.9-100	39.0 (37.6) 0-100.0	39.1 (36.5) 0-100.0	52.9 (33.3) 1.1-99.8
CD44 stroma + epithelium	27.0 (32.1) 0-98.6	25.8 (30.4) 0-99.1	32.4 (30.0) 0.2-96.7	21.2 (30.5) 0-91.3	25.3 (29.8) 0-92.5	27.4 (29.1) 0.4-96.7	28.3 (32.5) 0-98.6	22.9 (30.6) 0-99.1	35.5 (30.5) 0.2-96.7
CD24 stroma	6.9 (12.8) 0-63.5	8.4 (13.4) 0-83.6	12.2 (15.6) 0-73.0	5.4 (8.5) 0-37.1	5.7 (8.8) 0-46.2	10.9 (12.7) 0-42.3	7.2 (13.6) 0-63.5	9.1 (14.2) 0-83.6	13.1 (17.3) 0-73.0
CD24 epithelium	28.0 (22.6) 0.5-98.1	28.2 (23.1) 0-97.6	33.5 (23.0) 0.7-82.5	27.8 (23.3) 0.5-85.4	25.5 (20.8) 0-87.9	32.8 (23.1) 4.4-78.8	28.0 (22.5) 1.7-98.1	28.9 (23.6) 0-97.6	33.9 (23.3) 0.7-82.5
CD24 stroma + epithelium	17.4 (16.9) 0.2-91.6	18.3 (17.7) 0-93.2	22.8 (17.8) 0-74.1	16.3 (14.7) 0.2-64.4	16.4 (14.6) 0.3-63.2	20.9 (15.2) 0-49.7	17.7 (17.4) 0.2-91.6	18.7 (18.4) 0-93.2	23.9 (19.3) 0.8-74.1
ALDH1A1 stroma	11.1 (12.0) 0-85.0	10.4 (12.0) 0-67.2	17.3 (17.8) 0-64.6	7.1 (8.6) 0-34.5	12.4 (14.3) 0-67.2	12.9 (16.1) 0.5-64.6	11.9 (14.2) 0-85.0	9.9 (11.4) 0-56.8	20.0 (18.4) 0-64.6
ALDH1A1 epithelium	28.0 (21.1) 0.5-90.8	26.0 (19.3) 0.5-93.4	28.4 (17.1) 0.9-67.9	24.5 (26.2) 2.2-85.1	27.8 (19.6) 0.7-85.5	28.6 (18.8) 0.9-67.9	28.8 (20.0) 0.5-90.8	25.5 (19.3) 0.5-93.4	28.3 (16.2) 1.6-59.5
ALDH1A1 stroma + epithelium	21.0 (15.9) 0.4-88.5	19.5 (14.2) 1.0-70.1	23.9 (15.1) 1.1-56.8	17.4 (17.5) 1.8-62.3	21.7 (14.7) 1.0-70.1	21.7 (14.4) 1.1-55.4	21.7 (15.5) 0.4-88.5	18.9 (14.1) 1.2-67.0	25.1 (15.6) 1.3-56.8

Abbreviations: BBD, benign breast disease; SD, standard deviation.



Supplementary Figure 1. Associations of binary stem cell marker expression with subsequent breast cancer risk (ORs ad 95% CI).

Supplementary Table 2. Associations of stem cell markers in normal terminal duct-lobular units with subsequent breast cancer risk, adjusted for benign breast disease subtype (odds ratios and 95% confidence intervals)[a]

Stem cell marker expression	CD44		CD24		ALDH1A1	
	N cases/ controls	OR (95% CI)	N cases/ controls	OR (95% CI)	N cases/ controls	OR (95% CI)
Stroma						
Continuous, per 10% increase	94/394	0.95 (0.86; 1.04)	95/333	0.83 (0.67; 1.03)	90/350	0.87 (0.72; 1.07)
<10%	67/197	Ref	71/248	Ref	59/191	Ref
≥10%	27/127	0.53 (0.31; 0.93)	24/85	0.69 (0.39; 1.25)	31/124	0.74 (0.44; 1.26)
0-10%	67/197	Ref	71/249	Ref	59/191	Ref
>10-50%	13/76	0.40 (0.19; 0.83)	24/67	0.90 (0.49; 1.65)	29/114	0.81 (0.47; 1.38)
>50%	14/51	0.74(0.36; 1.52)	0/17	NE	2/10	0.27 (0.04; 1.74)
p-trend		0.25		0.04		0.12
Epithelium						
Continuous, per 10% increase	91/300	0.95 (0.88; 1.02)	92/312	0.95 (0.85; 1.07)	90/299	1.02 (0.90; 1.17)
<10%	30/93	Ref	26/63	Ref	26/67	Ref
≥10%	61/207	0.94 (0.55; 1.62)	66/249	0.62 (0.35; 1.11)	64/232	0.63 (0.35; 1.13)
0-10%	30/93	Ref	26/63	Ref	26/67	Ref
>10-50%	35/96	1.29 (0.70; 2.39)	48/188	0.63 (0.34; 1.15)	52/196	0.59 (0.32; 1.09)
>50%	26/111	0.66 (0.34; 1.27)	18/61	0.59 (0.27; 1.29)	12/36	0.85 (0.35; 2.04)
p-trend		0.11		0.28		0.67
Stroma + epithelial						
Continuous, per 10% change	94/325	0.95 (0.88; 1.04)	93/335	0.92 (0.79; 0.07)	96/326	0.98 (0.83; 1.16)
<10%	43/144	Ref	38/129	Ref	29/90	Ref
≥10%	51/181	0.96 (0.58; 1.61)	58/206	0.83 (0.50; 1.38)	67/236	0.77 (0.44; 1.35)
0-10%	44/144	Ref	38/129	Ref	29/91	Ref
>10-50%	28/99	1.05 (0.58; 1.90)	56/177	0.95 (0.56; 1.58)	64/220	0.80 (0.46; 1.40)
>50%	22/82	0.78 (0.41; 1.48)	2/29	0.18 (0.04; 0.85)	3/15	0.54 (0.13; 2.20)
p-trend		0.42		0.03		0.32

Abbreviations: CI, confidence interval; OR, odds ratio; NE, not estimable. [a]Adjusted for age (continuous), BMI (continuous), a family history of breast cancer (Yes/No), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/no hormones, postmenopausal/past hormones, postmenopausal/current hormones, postmenopausal/unknown hormone use status, unknown menopausal status), age at menarche (<12, 12, 13, >13, unknown), combined parity/age at first birth (parous with first birth before age 25, parous with first birth at or after age 25, nulliparous, unknown), alcohol use (none, >0-<5, ≥5 g/day), BBD subtype (non-proliferative, proliferative without atypia, and proliferative with atypia), and NHS cohort.

Supplementary Table 3. Associations of stem cell markers in normal terminal duct-lobular units with subsequent breast cancer risk, adjusted for percent breast density (odds ratios and 95% confidence intervals)[a]

Stem cell marker expression	CD44			CD24			ALDH1A1		
	N cases/ controls	OR (95% CI)	% mediated	N cases/ controls	OR (95% CI)	% mediated	N cases/ controls	OR (95% CI)	% mediated
Stroma									
Continuous, per 10% increase	20/137	0.82 (0.64; 1.06)	Not mediated	20/143	0.80 (0.48; 1.33)	6.4% (P=0.28)	19/138	0.76 (0.45; 1.31)	Not mediated
<10%	16/77	Ref		15/112	Ref		11/84	Ref	
≥10%	4/60	0.20 (0.05; 0.86)		5/31	1.09 (0.29; 4.12)		8/54	1.14 (0.38; 3.44)	
0-10%	16/77	Ref		15/112	Ref		11/84	Ref	
>10-50%	1/35	0.11 (0.01; 1.04)		5/26	1.39 (0.37; 5.27)		8/50	1.18 (0.39; 3.59)	
>50%	3/25	0.35 (0.06; 2.08)		0/54	NE		0/4	NE	
p-trend		0.12			0.40			0.91	
Epithelium									
Continuous, per 10% increase	19/129	0.85 (0.71; 1.01)	Not mediated	20/136	1.04 (0.78; 1.38)	Not mediated	20/130	1.32 (1.00; 1.75)	3.5% (P=0.29)
<10%	7/38	Ref		4/17	Ref		2/32	Ref	
≥10%	12/91	0.46 (0.14; 1.53)		16/119	0.27 (0.06; 1.29)		18/98	2.74 (0.51; 14.78)	
0-10%	7/38	Ref		4/17	Ref		2/32	Ref	
>10-50%	7/37	0.74 (0.19; 2.97)		11/97	0.23 (0.05; 1.15)		14/84	2.43 (0.44; 13.49)	
>50%	5/54	0.28 (0.06; 1.26)		5/22	0.54 (0.08; 3.78)		4/14	6.19 (0.67; 57.49)	
p-trend		0.09			0.80			0.11	
Stroma + epithelial									
Continuous, per 10% change	20/137	0.82 (0.66; 1.03)	Not mediated	21/144	0.94 (0.66; 1.34)	30.7% (P=0.21)	21/139	1.23 (0.83; 1.82)	7.7% (P=0.23)
<10%	12/53	Ref		7/46	Ref		3/39	Ref	
≥10%	8/84	0.32 (0.11; 1.01)		14/98	0.70 (0.22; 2.22)		18/100	1.97 (0.47; 8.27)	
0-10%	12/53	Ref		7/46	Ref		3/40	Ref	
>10-50%	4/44	0.35 (0.09; 1.38)		13/88	0.77 (0.24; 2.43)		18/93	2.12 (0.51; 8.93)	
>50%	4/40	0.30 (0.07; 1.40)		1/10	NE		0/6	NE	
p-trend		0.12			0.18			0.69	

Abbreviations: CI, confidence interval; OR, odds ratio; NE, not estimable. [a]Adjusted for age (continuous), BMI (continuous), a family history of breast cancer (Yes/No), menopausal status/postmenopausal hormone use (premenopausal, postmenopausal/no hormones, postmenopausal/past hormones, postmenopausal/current hormones, postmenopausal/unknown hormone use status, unknown menopausal status), age at menarche (<12, 12, 13, >13, unknown), combined parity/age at first birth (parous with first birth before age 25, parous with first birth at or after age 25, nulliparous, unknown), alcohol use (none, >0-<5, ≥5 g/day), percent breast density (continuous), and NHS cohort.