Differences in Risk Factors for Breast Cancer Molecular Subtypes in a Population-Based Study

Xiaohong R. Yang,[1] Mark E. Sherman,[1] David L. Rimm,[3] Jolanta Lissowska,[4] Louise A. Brinton,[1] Beata Peplonska,[6] Stephen M. Hewitt,[2] William F. Anderson,[1] Neonila Szeszenia-Dąbrowska,[6] Alicja Bardin-Mikolajczak,[4] Witold Zatonski,[4] Richard Cartun,[9] Daniza Mandich,[9] Grzegorz Rymkiewicz,[5] Marcin Ligaj,[5] Stanislaw Lukaszek,[8] Radzisaw Kordek,[7] and Montserrat García-Closas[1]

[1]Division of Cancer Epidemiology and Genetics, [2]Tissue Array Research Program, Laboratory of Pathology, Center For Cancer Research, National Cancer Institute, NIH, Department of Health and Human Services, Bethesda, Maryland; [3]Department of Pathology, Yale University School of Medicine, New Haven, Connecticut; [4]Department of Cancer Epidemiology and Prevention, and [5]Department of Pathology, Cancer Center and M. Sklodowska-Curie Institute of Oncology, Warsaw, Poland; [6]Department of Occupational and Environmental Epidemiology, Nofer Institute of Occupational Medicine, [7]Department of Pathology, Medical University of Łódź; [8]Department of Clinical Pathomorphology, Polish Mother's Memorial Hospital-Research Institute, Łódź, Poland; and [9]Department of Pathology, Hartford Hospital, Hartford, Connecticut

Abstract

Analysis of gene expression data suggests that breast cancers are divisible into molecular subtypes which have distinct clinical features. This study evaluates whether pathologic features and etiologic associations differ among molecular subtypes. We evaluated 804 women with invasive breast cancers and 2,502 controls participating in a Polish Breast Cancer Study. Immunohistochemical stains for estrogen receptor α, progesterone receptor, human epidermal growth factor receptors (HER2 and HER1), and cytokeratin 5 were used to classify cases into five molecular subtypes: luminal A, luminal B, HER2-expresing, basal-like, and unclassified. Relative risks were estimated using adjusted odds ratios and 95% confidence intervals. We observed that compared with the predominant luminal A tumors (69%), other subtypes were associated with unfavorable clinical features at diagnosis, especially HER2-expressing (8%) and basal-like (12%) tumors. Increasing body mass index significantly reduced the risk of luminal A tumors among premenopausal women (odds ratios, 0.71; 95% confidence intervals, 0.57-0.88 per five-unit increase), whereas it did not reduce risk for basal-like tumors (1.18; 0.86-1.64; $P_{heterogeneity}$ = 0.003). On the other hand, reduced risk associated with increasing age at menarche was stronger for basal-like (0.78; 0.68-0.89 per 2-year increase) than luminal A tumors (0.90; 0.95-1.08; $P_{heterogeneity}$ = 0.0009). Although family history increased risk for all subtypes (except for unclassified tumors), the magnitude of the relative risk was highest for basal-like tumors. Results from this study have shown that breast cancer risk factors may vary by molecular subtypes identified in expression studies, suggesting etiologic, in addition to clinical, heterogeneity of breast cancer. (Cancer Epidemiol Biomarkers Prev 2007;16(3):439–43)

Introduction

Gene expression profiling in tumor tissues suggests that breast cancers may be divided into subtypes consisting of two estrogen receptor (ER)–positive types (luminal A and luminal B) and three ER-negative types [human epidermal growth factor receptor 2 (HER2)–expressing, basal-like, and unclassified (''normal-like'')] with distinctive clinical outcomes (1, 2). Specifically, luminal A tumors, characterized by positive ER/progesterone receptor (PR) and negative HER2, show the most favorable clinical features among the five subtypes. Luminal B tumors express HER1 and HER2 in addition to ER/PR and show less favorable clinical outcomes compared with luminal A tumors. Basal-like tumors are characterized by the expression of cytokeratins 5/6 (CK5/6) and CK17 and are prevalent in patients with *BRCA1* mutations (2). Basal-like and HER2-overexpressing groups both are ER/PR-negative and have been associated with poor clinical features and survival. This classification has been highly consistent in independent studies using different array platforms, tumor sets, and statistical analyses. Studies using selected immunohistochemical stains have achieved similar stratifications of tumors according to clinical outcomes, suggesting that this molecular classification is robust (3, 4).

Data suggests that the molecular profiles in breast tumors are generally fixed at inception (5). Therefore, exposures that influence the risk of developing breast cancer might be related to the tumor molecular profiles that later affect the biology and clinical behavior of the tumors that arise. However, these molecular subtypes have only been evaluated in a single large population-based study (4), and the differences in etiology have not been investigated. Etiologic heterogeneity of breast cancer by hormone receptor status is supported by studies showing that age-specific incidence rates (6, 7) and specific risk factors vary by hormone receptor status. Specifically, nulliparity, late age at first birth, obesity among postmenopausal women, and early menarche have been more strongly linked to ER and/or PR positive than receptor negative tumors (8-12). Although ER and PR status figure prominently in the assignment of breast cancer subtypes, the molecular classification separates some tumors with identical ER or PR status, which may permit the identification of greater heterogeneity than would be evident by analysis of receptor status alone. For example, basal-like tumors are more common among African American than Caucasian women, suggesting that these tumors have a distinct etiology (4). A better understanding of the etiology of ER and PR negative tumors, which account for ~20% of breast cancers with known receptor status, is

Received 9/21/06; revised 12/26/06; accepted 1/11/07.

Grant support: In part by intramural funds from the National Cancer Institute.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Xiaohong Yang, Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Room 7014, 6120 Executive Boulevard, Bethesda, MD 20892-7236. Phone: 301-594-7804; Fax: 301-402-4489. E-mail: royang@mail.nih.gov

doi:10.1158/1055-9965.EPI-06-0806

particularly important because they include most clinically aggressive tumors. Given that increased understanding of breast cancer etiology may provide opportunities for improved prevention, we evaluated whether breast cancer risk factors vary according to molecular subtypes by analyzing data from a large population-based case-control study conducted in Poland.

Materials and Methods

Study Population. The study population has been previously described in detail (13). In brief, eligible subjects included women between the ages of 20 and 74 years who resided in Warsaw or Łódź in 2000 to 2003. Breast cancer cases with cytologically or pathologically confirmed incident *in situ* or invasive breast tumors were identified through a rapid identification system organized at five participating hospitals (~90% of cases) and cancer registries to ensure complete case ascertainment. Control subjects were randomly selected using a population-based database, frequency matched to cases on city and age in 5-year categories. A total of 2,386 cases (79% of those eligible) and 2,502 controls (69% of those eligible) agreed to participate in the study and provided informed consent as required by the National Cancer Institute and local institutional review boards. This report is based on all controls and a subset of 842 participating cases with invasive tumors arrayed on tissue microarray (TMA) blocks.

Information on breast cancer risk factors was assessed through a personal interview. Breast cancer risk factors evaluated in this analysis included well-established risk factors for breast cancer. Women were considered premenopausal if still menstruating at the time of interview, postmenopausal if periods had stopped, and unclear menopausal status if they began using hormone replacement therapy (HRT) before natural periods had stopped. Body mass index (BMI) was calculated using measured weight (kg) divided by standing height (m) squared. For ~5% of subjects with no measures of weight and height, BMI was calculated from self-reported information. History of a benign breast biopsy 1 year prior to date of diagnosis for cases, and date of interview for controls, was used as a surrogate for benign breast disease.

Pathology. Histopathologic features including diagnosis, grade, tumor size, and axillary lymph node metastases were assessed using clinical reports and independent evaluation by the study pathologist (M.E. Sherman). Routinely prepared formalin-fixed paraffin-embedded blocks of 842 invasive breast cancers were used to construct TMA blocks with 2-fold representation as 0.6 mm diameter cores per tumor (Beecher Instruments, Silver Spring, MD). Tissue sections of 5-μm thickness were placed on glass slides using a tape-transfer system (Instrumedics, Inc., Hackensack, NJ) with UV cross-linking, dipped in paraffin, and stored at room temperature under nitrogen to prevent oxidation-related loss of immunoreactivity prior to staining.

We did immunohistochemical stainings of TMA blocks for the markers required to recreate the molecular classification of tumors (3, 4). Staining was done with antigen retrieval prior to antibody incubation according to established protocols for ER-α (clone 6F11, 1:200; Novocastra, Newcastle upon Tyne, United Kingdom), PR (clone PgR636, 1:1,000; Dako), HER2 (polyclonal, 1:2,000; DakoCytomation, Glostrup, Denmark), HER1 (31G7, 1:500; Zymed, San Francisco, CA), and CK5 (XM26, 1:500; Novocastra). The study pathologist assessed the stains by digital image (Aperio, Vista, CA) and microscopic examination (M.E. Sherman). For tumors with two satisfactory cores (~80%), results were averaged; for the remainder, results were based on a single interpretable core. The percentage (0-100%) and intensity (0 = negative, 1 = weak, 2 = intermediate, and 3 = strong) of tumor cells stained were recorded for each marker. Stains for HER1 and CK5 were considered positive if any tumor cells were stained. ER and PR were classified as positive if the product of intensity and percentage was >10. HER2 was considered positive if intermediate or strong staining was identified in at least 20% of tumor cells as proposed previously (3). Among 842 tumors placed in the TMA blocks, 804 (95%) had interpretable staining results for all five markers evaluated. The subsequent analyses were based on data obtained from these 804 tumors. There were no significant differences in age or frequency of any of the established breast cancer risk or protective factors between the cases represented in the TMA blocks that had interpretable staining results for all five markers (n = 804) and the remaining cases in the study.

Tumor subtypes were defined as luminal A (ER+ and/or PR+, HER2−, any CK5/HER1), luminal B (ER+ and/or PR+, HER2+, any CK5/HER1), HER2-expressing (ER−, PR−, HER2+, any CK5/HER1), basal-like (ER−, PR−, Her2−, CK5+, and/or HER1+), and unclassified or normal-like (negative for all five markers; ref. 4).

Statistical Analysis. Differences between breast cancer subtypes with regard to clinicopathologic characteristics and risk factors were examined using one-way ANOVA for continuous variables and χ^2 tests for the categorical variables.

Table 1. Tumor characteristics for different breast cancer subtypes among 804 women with invasive breast cancer participating in the Polish Breast Cancer Study

Characteristics	Luminal A	Luminal B	HER2-expressing	Basal-like	Unclassified	P*
No. (% of total)	552 (69)	48 (6)	61 (8)	95 (12)	48 (6)	
Histology, no. (%)						
Ductal	308 (56)	36 (75)	54 (89)	63 (80)	36 (57)	
Lobular	128 (23)	5 (10)	3 (5)	8 (10)	10 (16)	<0.0001
Mixed and other	66 (21)	7 (14)	4 (7)	8 (10)	17 (27)	Not determined
Tumor grade, no. (%)						
Well differentiated	149 (27)	1 (2)	0 (0)	2 (2)	7 (15)	
Moderately differentiated	345 (62)	30 (63)	22 (36)	25 (26)	26 (54)	
Poorly differentiated	58 (11)	17 (35)	39 (64)	68 (72)	15 (31)	<0.0001 †
Tumor size (cm), no. (%)						
≤2	332 (60)	16 (33)	19 (32)	36 (38)	14 (29)	
>2	218 (40)	32 (67)	41 (68)	59 (62)	34 (71)	<0.0001
Axillary node metastasis, no. (%)						
Negative	330 (61)	24 (52)	26 (43)	54 (57)	26 (54)	
Positive	210 (39)	22 (48)	35 (57)	40 (43)	22 (46)	0.06

NOTE: Tumor subtypes were defined as luminal A (ER+ and/or PR+, HER2−), luminal B (ER+ and/or PR+, HER2+), HER2 (HER2+, ER−, PR−), basal-like (ER−, PR−, HER2−, CK5+, and/or HER1+), and unclassified or normal-like (negative for all five markers; ref. 4).
*P values were obtained by comparing the five tumor subtypes using χ^2 test for differences in frequencies.
†Well and moderately differentiated tumors were combined into one category for P value calculation.

Table 2. Distribution of established breast cancer etiologic factors among 804 breast cancer cases and 2,502 controls participating in the Polish Breast Cancer Study

Subject characteristics	Controls* (N = 2,502)	Breast cancer cases*,†					
		Luminal A (n = 552)	Luminal B (n = 48)	HER2-expressing (n = 61)	Basal-like (n = 95)	Unclassified (n = 48)	P ‡
Age, mean (SD)	55.9 (10.1)	56.3 (9.9)	56.3 (9.4)	56.0 (10.6)	53.7 (10.9)	57.3 (9.9)	0.27
Education, no. (%)							
High school or lower	1,900 (76)	343 (63)	35 (73)	44 (72)	65 (69)	34 (71)	
College or higher	586 (24)	204 (37)	13 (27)	17 (28)	29 (31)	14 (29)	0.25
Age at menarche, mean (SD)	13.7 (1.7)	13.5 (1.7)	13.4 (1.9)	13.7 (1.9)	13.0 (1.6)	13.7 (1.3)	0.003
Age at first full term birth, mean (SD)	23.6 (4.2)	24.6 (4.5)	23.8 (4.1)	23.5 (3.7)	23.7 (4.0)	24.2 (3.9)	0.002
Number of full-term births, no. (%)							
0	281 (11)	99 (18)	7 (15)	11 (18)	9 (9)	6 (13)	
1	746 (30)	185 (33)	18 (37)	19 (31)	36 (38)	13 (27)	0.36
≥2	1,475 (59)	268 (49)	23 (48)	31 (51)	50 (53)	29 (60)	0.29
Menopausal status, no. (%)							
Premenopausal	715 (29)	144 (26)	12 (25)	12 (20)	36 (38)	13 (27)	
Postmenopausal	1,682 (67)	366 (66)	33 (69)	47 (77)	56 (59)	33 (69)	0.13
Unclear	103 (4)	42 (8)	3 (6)	2 (3)	3 (3)	2 (4)	Not determined
Age at menopause, mean (SD)	49.2 (5.0)	49.8 (4.4)	49.6 (4.6)	49.3 (4.7)	49.4 (4.0)	51.2 (4.3)	0.07
BMI among premenopausal women, mean (SD)	26.4 (5.1)	24.8 (4.0)	25.6 (5.8)	23.5 (4.1)	27.1 (4.9)	27.2 (4.8)	0.003
BMI among postmenopausal women, mean (SD)	28.6 (5.4)	28.3 (5.6)	28.7 (5.7)	28.2 (5.0)	27.8 (5.4)	28.9 (6.6)	0.80
Oral HRT use among postmenopausal women, no. (%)							
Never user	1,323 (83)	260 (76)	28 (88)	38 (83)	42 (81)	20 (65)	
Current/recent use of combined HRT	93 (6)	45 (13)	1 (3)	4 (9)	4 (7)	5 (16)	0.25
Past use of combined HRT	75 (5)	19 (5)	0 (0)	2 (4)	3 (6)	2 (6)	Not determined
Use of estrogen or progesterone alone	104 (6)	20 (6)	3 (9)	2 (4)	3 (6)	4 (13)	Not determined
Breast cancer in first-degree relatives							
No	2,356 (94)	496 (90)	43 (90)	53 (87)	80 (84)	45 (94)	
Yes	146 (6)	55 (10)	5 (10)	8 (13)	15 (16)	3 (6)	0.37
Previous benign breast disease, no. (%)							
No	2,323 (94)	476 (89)	43 (91)	55 (92)	87 (95)	43 (93)	
Yes	157 (6)	61 (11)	4 (9)	5 (8)	5 (5)	3 (7)	0.39

*Numbers for some factors might be less than the total numbers due to missing data.
†Molecular subtypes were defined as luminal A (ER+ and/or PR+, HER2−), luminal B (ER+ and/or PR+, HER2+), HER2 (HER2+, ER−, PR−), basal-like (ER−, PR−, HER2−, CK5+, and/or HER1+), and unclassified or normal-like (negative for all five markers; ref. 4).
‡ P values were calculated to compare the five tumor subgroups, using ANOVA for differences in mean, and χ^2 or Fisher exact tests for differences in frequencies.

The Fisher's exact test was used when expected cell counts were less than 5. Multivariate polychotomous logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (95% CI) for the association between exposures and each breast cancer subtype with molecular subtypes as outcome variables and exposures as explanatory variables. The final models included the following variables: age, study site, education level, age at menarche, age at menopause, menopausal status, number of full-term births, age at first full-term birth, current/recent oral HRT use among postmenopausal women, previous breast disease (history of a benign breast biopsy 1 year prior to date of diagnosis for cases and date of interview for controls), mammographic screening (ever/never), family history of breast cancer among first-degree relatives, and BMI. BMI among premenopausal and postmenopausal women were analyzed separately as two variables in the models to account for the effect modification by menopausal status. Factors such as alcohol consumption (ever/never and duration), oral contraceptive use (ever/never, not common in the Polish population), and smoking were not significantly associated with breast cancer risk or molecular subtypes and thus were not included in the final model. P values to test for heterogeneity of exposure-disease ORs between different tumor subtypes were obtained using logistic regression analyses restricted to cases, with tumor subtypes as the outcome and exposures as the explanatory variables. To disentangle the associations of correlated tumor characteristics and molecular subtypes on exposures, we fitted regression models with the exposure of interest as the outcome and tumor characteristics as the explanatory variables. All statistical analyses were done using the SAS (version 8.0, SAS Institute, Inc., Cary, NC) software.

Results

Characteristics of Participants. There were no significant differences in the age or frequency of any of the established breast cancer risk or protective factors between the cases represented in the TMA blocks that had interpretable staining results for all five markers (n = 804) and the remaining cases in the study (data not shown). Because of limited availability of tissue, cases with small tumors were underrepresented in the TMA blocks (7% of cases with tissue samples in the TMA blocks had tumors ≤1 cm compared with 20% of cases not included in the TMA blocks; P < 0.0001). Other tumor characteristics (i.e., histology, grade, nodal status, and hormone receptor status) were similar after adjusting for size.

Pathologic Features of Breast Cancer Subtypes. Table 1 shows the distribution of tumor characteristics among the five

molecular subtypes. Luminal A tumors (n = 552) comprised 69% of tumors followed by less common subtypes: basal-like (12%, n = 95), HER2-expressing (8%, n = 61), luminal B (6%, n = 48), and unclassified (6%, n = 48). Most (75-89%) tumors in luminal B, HER2-expressing and basal-like categories were ductal carcinomas. In contrast, luminal A tumors included 56% ductal and 23% lobular tumors. Luminal A tumors included the lowest frequency of poorly differentiated tumors, whereas HER2-expressing and basal-like tumors were associated with the highest frequency (P < 0.0001). Luminal A tumors also included a greater frequency of small tumors (≤2 cm) compared with other subtypes (P < 0.0001); however, the frequency of axillary node metastases did not differ significantly by tumor subtypes (Table 1).

Distribution of Risk Factors by Tumor Subtypes. We observed significant differences by tumor subtype for the distribution of age at menarche (P = 0.003), age at first full-term birth (P = 0.002), and BMI among premenopausal women (P = 0.003; Table 2). Basal-like tumors were related to the earliest age at menarche and highest BMI among premenopausal women, whereas luminal A tumors were associated with the oldest mean age at first full-term birth. Frequencies for other exposures did not differ significantly by tumor subtypes, although some comparisons (e.g., HRT use) were limited by small numbers.

Determinants of Cancer Risk by Tumor Subtypes. Table 3 shows the estimated ORs and 95% CIs for the associations between exposures and tumor subtypes based on case-control analyses, as well as the P values for heterogeneity of exposures across tumor subtypes using the luminal A tumors as the reference group. Increasing age at menarche was associated with lower risk of breast cancer for basal-like than other tumors ($P_{heterogeneity}$ = 0.0009 compared with luminal A). Increased BMI among premenopausal women reduced the risk of luminal A, B, and HER2-expressing tumors, whereas no protective association was seen against basal-like tumors ($P_{heterogeneity}$ = 0.003) or unclassified tumors ($P_{heterogeneity}$ = 0.12). Differential associations of age at menarche and BMI among premenopausal women for luminal A and basal-like groups remained significant after adjustment of other tumor variables including tumor size, grade, histologic type, and axillary nodes (P = 0.01 for age at menarche and P = 0.02 for BMI). Associations for parity and family history of breast cancer did not differ significantly for subtypes, although the strongest reverse risk from parity was seen for luminal A tumors, whereas the greatest increase in risk associated with family history was found for basal-like tumors based on limited numbers. Associations with age at first full-term birth and BMI among postmenopausal women differed minimally across molecular subtypes. Age at menopause for women with premenopausal hysterectomies but no oophorectomy was defined as age at surgery, which would tend to underestimate the true age at menopause for these women. We did a separate analysis by restricting the regression analysis to women with natural menopause. Estimates for age at menopause or other variables in the model did not vary appreciably from those obtained when all postmenopausal women were included (data not shown).

Discussion

Analysis of data from this population-based study showed differences in risk or protective factors by tumor subtypes, supporting the view that the molecular classification might be relevant for understanding tumor etiology. In addition, we confirmed previously reported differences in tumor pathology and patient characteristics by breast cancer subtypes (1-4, 14).

Consistent with prior reports, we found that luminal A breast cancers were numerically predominant and differ in pathologic characteristics from non–luminal A tumors. Specifically, luminal A tumors included a higher percentage of lobular carcinomas, the lowest frequency of poorly differentiated carcinomas, and the highest frequency of small tumors in this relatively infrequently screened population (54% of control women reported ever having had a screening mammogram). In contrast, HER2-expressing and basal-like tumors showed the highest frequency of poorly differentiated carcinomas. These data are consistent with previous analyses showing that basal-like tumors are associated with poor

Table 3. Association between selected breast cancer risk or protective factors and breast cancer subtypes among 804 cases and 2,502 controls participating in the Polish Breast Cancer Study

Subject characteristics	Luminal A (n = 552)	Luminal B (n = 48)		HER2-expressing (n = 61)		Basal-like (n = 95)		Unclassified (n = 48)	
	OR* (95% CI)	OR* (95% CI)	P †	OR* (95% CI)	P †	OR* (95% CI)	P †	OR* (95% CI)	P †
Age at menarche (per 2 y increase)	0.90 (0.95-1.08)	0.98 (0.75-1.28)	0.99	1.14 (0.86-1.50)	0.26	0.78 (0.68-0.89)	0.0009	0.92 (0.72-1.18)	0.61
Number of full-term births ‡									
1	0.50 (0.24-1.06)	0.91 (0.09-8.83)	0.87	1.50 (0.19-11.77)	0.31	2.36 (0.43-13.01)	0.14	0.57 (0.06-5.83)	0.72
≥2	0.42 (0.21-0.84)	0.56 (0.07-4.56)	0.94	1.15 (0.17-7.63)	0.29	1.80 (0.37-8.85)	0.20	0.71 (0.09-5.71)	0.44
Age at first full-term birth (per 5-y increase)	1.08 (0.95-1.23)	1.04 (0.70-1.55)	0.95	0.87 (0.60-1.05)	0.27	0.95 (0.71-1.27)	0.47	1.06 (0.71-1.58)	0.71
Age at menopause (per 5-y increase)	1.13 (1.01-1.28)	1.10 (0.78-1.57)	0.93	1.09 (0.81-1.46)	0.82	1.02 (0.82-1.28)	0.38	1.46 (0.99-2.15)	0.19
BMI in premenopausal women (per five units)	0.71 (0.57-0.88)	0.88 (0.48-1.60)	0.37	0.53 (0.25-1.12)	0.38	1.18 (0.86-1.64)	0.003	1.17 (0.68-2.00)	0.12
BMI in postmenopausal women (per five units)	1.00 (0.90-1.12)	1.02 (0.74-1.42)	0.91	0.97 (0.73-1.28)	0.99	0.87 (0.66-1.14)	0.37	1.08 (0.78-1.48)	0.85
Family history §	1.72 (1.21-2.45)	2.31 (0.88-6.08)	0.61	2.35 (1.03-5.38)	0.57	3.17 (1.69-5.92)	0.11	0.94 (0.28-3.16)	0.49

NOTE: Molecular subtypes were defined as luminal A (ER+ and/or PR+, HER2−), luminal B (ER+ and/or PR+, HER2+), HER2 (HER2+, ER−, PR−), basal-like (ER−, PR−, HER2−, CK5+, and/or HER1+), and unclassified or normal-like (negative for all five markers; ref. 4).

*ORs and 95% CI were obtained from case-control comparisons using polytomous logistic regression analyses adjusting for age, study site, education level, menopausal status, oral HRT use among postmenopausal women, previous benign breast disease, mammogram screening, and all factors shown in the table.

†P values were obtained from comparisons of each tumor subtype to luminal A group using unconditional logistic regression analysis adjusting for all risk factors mentioned in the first footnote.

‡Compared with nulliparous women.

§Family history of breast cancer among first-degree relatives; compared with no family history.

survival after adjusting for nodal status and other features (2-4, 14), and analyses indicating that HER2 gene amplification, which generally corresponds to strong immunohistochemical expression, is a poor prognostic factor (15).

Luminal A tumors were associated with most exposures that have been consistently related to overall breast cancer risk. Differences were observed between luminal A and other subtypes. In particular, the protection from late age at menarche was strongest for basal-like tumors, however, elevated BMI among premenopausal women did not reduce the risk of this tumor subtype. The contrasting associations for obesity in premenopausal women with basal-like as compared with luminal A tumors observed in this study are consistent with our current understanding of hormonal carcinogenesis, given that the former are ER- and PR-positive and the latter ER- and PR-negative. Based on limited data, obesity in premenopausal women reduces estrogen exposure secondary to anovulation (16), which could account for the protective effect of obesity for receptor-positive luminal A tumors compared with receptor-negative basal-like tumors. The differential associations of BMI in premenopausal women and age at menarche among breast cancer subtypes raise public health concerns, given data suggesting that BMI is increasing among girls and young women in Western countries and that age at menarche has declined. These secular trends have been seen more strikingly among African Americans in the U.S. (17-19), which is notable in light of the higher frequency of basal-like tumors in this racial group (4).

The inverse association of parity was not seen for basal-like tumors, whereas the risk associated with family history was greatest for this tumor type. In combination with some studies that have found younger age at diagnosis for basal-like tumors and associations with germ line *BRCA1/2* mutations, our data suggest a particularly strong influence of genetics in the etiology of these tumors (2, 20). Our results are also in line with previous demonstrations that breast cancer risk among BRCA1 carriers was significantly associated with early age at menarche (21). Similarly, an early first full-term birth and obesity did not confer protection for BRCA1 carriers (22) and weight gain in early years (≤age 30) even increased the risk of breast cancer among BRCA1 carriers (23, 24).

Although based on small numbers, our analysis suggests that HER2-expressing tumors may have different risk factor associations compared with luminal A tumors. Specifically, relative risk estimates for older age at menarche and increasing parity did not show protection against HER2-expressing versus luminal A carcinoma subtypes. Re-examination of these associations in larger data sets is needed to confirm these findings. Differential associations by tumor subtype related to patterns of breast development, features of the premenopausal hormonal milieu, or genetic background may account for the differences observed in these associations.

Our study is the largest population-based study to evaluate molecular subtypes of breast cancer, with a comprehensive collection of breast cancer risk factors. This is a unique feature because most previous studies of molecular subtypes were based on small sets of clinical samples without collection of epidemiologic information. In addition, we used TMA technology to make assay conditions for measuring marker expression more standardized. Moreover, we applied a semiquantitative scoring system versus a positive-negative dichotomization to minimize misclassification. However, evaluation of TMA blocks may lead to misclassification of marker expression for tumors with regional differences in marker expression levels, which would tend to dilute associations with exposures. In addition, despite being one of the largest studies of molecular subtypes, our analysis lacked power to precisely assess associations for uncommon tumor types or exposures. Nonetheless, our data supports the hypothesis that breast cancer risk factors vary by molecular-based breast cancer subtypes. In particular, luminal A and basal-like tumors seem to have distinct sets of risk factors, suggesting that these two molecular subtypes may be etiologically different. Moreover, improved understanding of non–luminal A breast cancer subtypes associated with poorer prognosis may have important public health relevance.

Acknowledgments

We thank Anita Soni, Elena Adrianza (Westat, Rockville, MD) for their work on study management; Pei Chao (IMS, Silver Spring, MD) for her work on data and sample management; Lori Charette (TMA core facility, Yale University, New Haven, CT) for her technical support on TMA construction; the physicians, pathologists, and nurses from participating centers in Poland as well as interviewers and study participants for their efforts during field-work.

References

1. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature 2000;406:747–52.
2. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003;100:8418–23.
3. Nielsen TO, Hsu FD, Jensen K, et al. Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. Clin Cancer Res 2004;10:5367–74.
4. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA 2006;295:2492–502.
5. Lacroix M, Toillon RA, Leclercq G. Stable 'portrait' of breast tumors during progression: data from biology, pathology and genetics. Endocr Relat Cancer 2004;11:497–522.
6. Yasui Y, Potter JD. The shape of age-incidence curves of female breast cancer by hormone-receptor status. Cancer Causes Control 1999;10:431–7.
7. Anderson WF, Chatterjee N, Ershler WB, Brawley OW. Estrogen receptor breast cancer phenotypes in the Surveillance, Epidemiology, and End Results database. Breast Cancer Res Treat 2002;76:27–36.
8. Althuis MD, Fergenbaum JH, Garcia-Closas M, Brinton LA, Madigan MP, Sherman ME. Etiology of hormone receptor-defined breast cancer: a systematic review of the literature. Cancer Epidemiol Biomarkers Prev 2004;13:1558–68.
9. Colditz GA, Rosner BA, Chen WY, Holmes MD, Hankinson SE. Risk factors for breast cancer according to estrogen and progesterone receptor status. J Natl Cancer Inst 2004;96:218–28.
10. Cotterchio M, Kreiger N, Theis B, Sloan M, Bahl S. Hormonal factors and the risk of breast cancer according to estrogen- and progesterone-receptor subgroup. Cancer Epidemiol Biomarkers Prev 2003;12:1053–60.
11. Huang WY, Newman B, Millikan RC, Schell MJ, Hulka BS, Moorman PG. Hormone-related factors and risk of breast cancer in relation to estrogen receptor and progesterone receptor status. Am J Epidemiol 2000;151:703–14.
12. Rusiecki JA, Holford TR, Zahm SH, Zheng T. Breast cancer risk factors according to joint estrogen receptor and progesterone receptor status. Cancer Detect Prev 2005;29:419–26.
13. Garcia-Closas M, Brinton LA, Lissowska J, et al. Established breast cancer risk factors by clinically important tumour characteristics. Br J Cancer 2006; 95:123–5.
14. van de RM, Perou CM, Tibshirani R, et al. Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. Am J Pathol 2002;161:1991–6.
15. Menard S, Fortis S, Castiglioni F, Agresti R, Balsari A. HER2 as a prognostic factor in breast cancer. Oncology 2001;61:67–72.
16. Cleary MP, Maihle NJ. The role of body mass index in the relative risk of developing premenopausal versus postmenopausal breast cancer. Proc Soc Exp Biol Med 1997;216:28–43.
17. Biro FM, Huang B, Crawford PB, et al. Pubertal correlates in black and white girls. J Pediatr 2006;148:234–40.
18. Demerath EW, Towne B, Chumlea WC, et al. Recent decline in age at menarche: the Fels Longitudinal Study. Am J Hum Biol 2004;16:453–7.
19. Himes JH. Examining the evidence for recent secular changes in the timing of puberty in US children in light of increases in the prevalence of obesity. Mol Cell Endocrinol 2006;254–5:13–21.
20. Foulkes WD, Stefansson IM, Chappuis PO, et al. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. J Natl Cancer Inst 2003;95: 1482–5.
21. Kotsopoulos J, Lubinski J, Lynch HT, et al. Age at menarche and the risk of breast cancer in BRCA1 and BRCA2 mutation carriers. Cancer Causes Control 2005;16:667–74.
22. Jernstrom H, Lerman C, Ghadirian P, et al. Pregnancy and risk of early breast cancer in carriers of BRCA1 and BRCA2. Lancet 1999;354:1846–50.
23. Kotsopoulos J, Olopado OI, Ghadirian P, et al. Changes in body weight and the risk of breast cancer in BRCA1 and BRCA2 mutation carriers. Breast Cancer Res 2005;7:R833–43.
24. Nkondjock A, Robidoux A, Paredes Y, Narod SA, Ghadirian P. Diet, lifestyle and BRCA-related breast cancer risk among French-Canadians. Breast Cancer Res Treat 2006;98:285–94.

Differences in Risk Factors for Breast Cancer Molecular Subtypes in a Population-Based Study

Xiaohong R. Yang, Mark E. Sherman, David L. Rimm, et al.

Cancer Epidemiol Biomarkers Prev 2007;16:439-443.