

OPEN

# Attribution to Heterogeneous Risk Factors for Breast Cancer Subtypes Based on Hormone Receptor and Human Epidermal Growth Factor 2 Receptor Expression in Korea

Boyoung Park, MD, PhD, Ji-Yeob Choi, PhD, Ho Kyung Sung, MD, Choonghyun Ahn, MD, Yunji Hwang, MS, Jieun Jang, BS, Juyeon Lee, BS, Heewon Kim, BS, Hai-Rim Shin, MD, PhD, Sohee Park, PhD, Wonshik Han, MD, PhD, Dong-Young Noh, MD, PhD, Keun-Young Yoo, MD, PhD, Daehee Kang, MD, PhD, and Sue K. Park, MD, PhD

**Abstract:** We conducted a heterogeneous risk assessment of breast cancer based on the hormone receptor (HR) and human epidermal growth factor receptor 2 (HER2) calculating the risks and population-based attributable fractions (PAFs) for modifiable and nonmodifiable factors.

Using matched case-control study design from the Seoul Breast Cancer Study and the national prevalence of exposure, the risks and PAFs for modifiable and nonmodifiable factors were estimated for total breast cancers and subtypes.

The attribution to modifiable factors was different for each subtype (luminal A, PAF = 61.4% [95% confidence interval, CI = 54.3%–69.8%]; luminal B, 21.4% [95% CI = 18.6–24.9%]; HER2-overexpression, 59.4% [95% CI = 47.8%–74.3%], and triple negative tumors [TNs], 27.1% [95% CI = 22.9%–32.4%]), and the attribution to the modifiable factors for the luminal A and HER2-overexpression subtypes was higher than that of the luminal B and TN subtypes ( $P$  heterogeneity  $\leq 0.001$ ). The contribution of modifiable reproductive factors to luminal A type in premenopausal women was higher than that of the other subtypes (18.2% for luminal A; 3.1%, 8.1%, and –3.1% for luminal B, HER2-overexpression, and TN subtypes, respectively;  $P$  heterogeneity  $\leq 0.001$ ). Physical activity had the highest impact preventing 32.6% of luminal A, 14.5% of luminal B, 38.0% of

HER2-overexpression, and 26.9% of TN subtypes ( $P$  heterogeneity = 0.014). Total reproductive factors were also heterogeneously attributed to each breast cancer subtype (luminal A, 65.4%; luminal B, 24.1%; HER2-overexpression, 57.9%, and TN subtypes, –3.1%;  $P$  heterogeneity  $\leq 0.001$ ).

Each pathological subtype of breast cancer by HRs and HER2 status may be associated with heterogeneous risk factors and their attributable risk, suggesting a different etiology. The luminal B and TN subtypes seemed to be less preventable despite intervention for alleged risk factors, even though physical activity had a high preventable potential against breast cancer.

(*Medicine* 95(14):e3063)

**Abbreviations:** BMI = body mass index, CI = confidence interval, ER = estrogen receptor, HER2 = human epidermal growth factor receptor 2, HR = hormone receptor, HRT = hormone replacement therapy, IARC = International Agency for Research on Cancer, KNHANES = Korea National Health and Nutrition Examination Surveys, OR = odds ratio, PAF = population-based attributable fractions, PR = progesterone receptor, RRs = relative risks, SeBCS = Seoul Breast Cancer Study, TN = triple negative.

Editor: Ching-Sheng Hsu.

Received: June 29, 2015; revised: January 30, 2016; accepted: February 20, 2016.

From the Graduate School of Cancer Science and Policy, National Cancer Center (BP); Department of Preventive Medicine, Seoul National University College of Medicine (J-YC, HKS, CA, YH, JJ, JL, HK, K-YY, DK, SKP); Cancer Research Center (J-YC, CA, YH, JJ, JL, HK, DK, SKP); Department of Biomedical Science, Seoul National University College of Medicine, Seoul, Republic of Korea (J-YC, HKS, CA, YH, JJ, JL, HK, DK, SKP); Western Pacific Regional Office, World Health Organization, Manila, Philippines (H-RS); Department of Epidemiology and Health Promotion, Graduate School of Public Health, Yonsei University (SP); and Department of Surgery, Seoul National University Hospital and College of Medicine (WH-D-YN), Seoul, Korea.

Correspondence: Sue K. Park, Department of Biomedical Science, Seoul National University Graduate School, Department of Preventive Medicine, Seoul National University College of Medicine, 103 Daehakro, Chongno-gu, Seoul 110–799, Korea (e-mail: suepark@snu.ac.kr).

This study was supported by a grant from the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (1420190). The funding source played no role in the study design, data collection and analysis, decision to publish, or manuscript preparation.

The authors report no conflicts of interest.

Supplemental Digital Content is available for this article.

Copyright © 2016 Wolters Kluwer Health, Inc. All rights reserved.

This is an open access article distributed under the Creative Commons Attribution-NonCommercial License, where it is permissible to download, share and reproduce the work in any medium, provided it is properly cited.

The work cannot be used commercially.

ISSN: 0025-7974

DOI: 10.1097/MD.0000000000003063

## INTRODUCTION

Breast cancer is the most common cancer among females, accounting for 22.9% of all cancers.<sup>1</sup> Although several risk factors have been identified, reproductive factors including age at menarche and menopause and family history of breast cancer are not amendable; however, lifestyle and behavioral risk factors such as physical inactivity, alcohol consumption, obesity, or exogenous hormone use could be modified to reduce the risk of breast cancer.<sup>2</sup> Previous studies have shown that population attributable fractions (PAFs), the proportion of cases prevented by removing risk factors from a population,<sup>3</sup> of several modifiable and non-modifiable risk factors ranged from 30% to >60% of all breast cancer cases.<sup>4–6</sup> These results on both the magnitude of risk and PAFs have been derived from epidemiological studies that considered breast cancer as a single disease despite the clinical, pathological, and molecular heterogeneity of tumor subtypes.<sup>7</sup>

A method using immunohistochemistry classifies breast cancer according to estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2),<sup>8</sup> and several studies have identified that risk factors were differently associated with breast cancer according to tumor subtypes based on receptor status (ER, PR, and/or HER2).<sup>7,9,10</sup> However, few studies investigated the PAFs

according to subtype and showed the heterogeneous effect of removing risk factors based on only ER and PR status, in which an intervention for modifiable factors was less relevant for hormone receptor (HR)–negative breast cancers.<sup>5</sup> Although HER2 is important to determine the biological characteristics of the tumor and the treatment,<sup>11</sup> estimation of the heterogeneity of PAFs based on both HR and HER2 status was rare. We conducted a risk assessment of breast cancer in Korea by calculating risks measured by odds ratios (ORs) and PAFs for modifiable and nonmodifiable risk factors based on HR and HER2 status and assessed the heterogeneity in risk factors and PAFs according to tumor subtype.

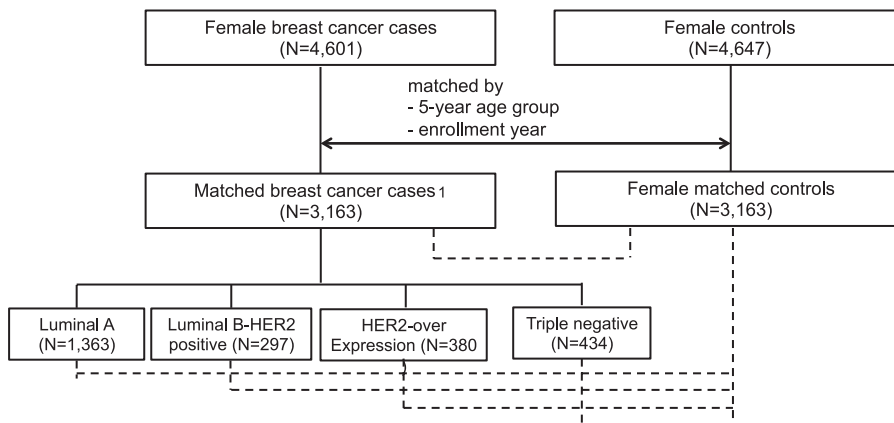
**METHODS**

The risk factors and their quantitative risk were assessed using data from the Seoul Breast Cancer Study (SeBCS), which is the largest multicenter-based case-control study of breast cancer in Korea. Details of the study have been previously published.<sup>12</sup> The cases (N = 4601) consisted of women diagnosed with breast cancer admitted to 3 university hospitals in Seoul from 2001 to 2007, which accounts for about 15% to 18% of the total treated breast cancer patients in Korea. The controls (N = 4647) were recruited from health examinees without cancers in the same hospitals as the cases or in community health screening programs in urban areas. The SeBCS was approved by the Ethics Committees of all participating hospitals. All participants gave written informed consent at the time of enrollment, and face-to-face interviews with trained interviewers were conducted using a structured questionnaire. We constructed 3163 sets of cases and controls matched by 5-year age group and enrollment year. Among 3163 cases, 2474 cases had information on all 3 receptors from pathology reports. Breast cancer patients from the SeBCS included nonclassified types (21.8%) by hormone and HER2 receptors. After exclusion of the nonclassified types, the distribution of each subtype was luminal A type (ER+ regardless of PR status and HER2–), 55.1% (N = 1363); luminal B-HER2 positive type (ER+ regardless of PR status and HER2+; luminal B type), 12.0% (N = 297); HER2-overexpression (ER–, PR–, and HER2+), 15.4% (N = 380), and triple negative (ER–, PR–, and HER2–; TN), 17.5% (N = 434) (Figure 1). The distribution of the 4 major subtypes in the SeBCS was comparable with the

distribution in the data of the Korean Breast Cancer Society Registry, in which the percentage of each subtype was 54.9%, 11.4%, 12.1%, and 21.6%, respectively.<sup>13</sup>

Family history of breast cancer (defined as family history in first degree relatives), age at menarche, age at menopause, age at first birth, total period of breastfeeding, oral contraceptive use (never and ever), body mass index (BMI) after menopause, and exercise (defined as regular leisure-time physical activity ≥1 hour per week) were selected as the significant risk factors in a backward multiple logistic regression model. Alcohol consumption (never and ever) was additionally included in a multivariate model, although it was not selected in the backward multiple logistic regression model because it was suggested as a proven breast carcinogenic agent in International Agency for Research on Cancer (IARC) criteria<sup>14,15</sup> and the prevalence of alcohol drinkers in Korea is among the highest in the world<sup>16</sup> with annually increased number of Korean women who drink.<sup>17</sup>

Although hormone replacement therapy (HRT) use is a known breast carcinogen with sufficient evidence according to the IARC criteria,<sup>14,15</sup> we did not include HRT use in the multivariate model because the point estimate of the OR for breast cancer was 1.0 in the univariate model, and its use is not common in Korea (<5% in postmenopausal women) with unknown components in early 1990s. Using these 9 variables, quantitative risks for total breast cancer cases and 4 subtypes were assessed by multivariable logistic regression adjusting for all variables. The 3163 controls were initially matched to all breast cancer cases and used as a comparative group not only for all breast cancers but also for the 4 subtypes. We also adjusted for age and enrollment year. We divided the 9 factors as modifiable factors and nonmodifiable factors. Nonmodifiable factors included family history of breast cancer, age at menarche, and age at menopause (only for postmenopausal women), and the modifiable factors included age at first birth, total period of breastfeeding and oral contraceptive use for premenopausal women, BMI for postmenopausal women, and exercise and alcohol drinking for all women. Age at first birth was included as modifiable factors, considering the effect of the population control policies through national birth control programs on the rapid decline of the fertility rate in Korea.<sup>22</sup> The distribution of selected factors in patients with different subtypes of breast cancer and controls is presented in Table 1.



**FIGURE 1.** Flow chart of case–control matching process and analysis scheme. A total of 3163 matched breast cancer cases included 689 breast cancer patients with nonclassified types by hormone and HER2 receptors. Dotted line represents comparison groups.

**TABLE 1.** Distribution of Major Risk Factors for Breast Cancer Regarding Overall Breast Cancer and Subtypes Classified by HR and HER2

	Controls (N = 3163)	Total cancer (N = 3163)*	Luminal A (N = 1,363)	Luminal B–HER2 Positive (N = 297)	HER2-overexpres- sion (N = 380)	Triple Negative (N = 434)
	%	%	%	%	%	%
Family history of breast cancer						
No	97.3	95.0	94.5	95.3	96.3	94.9
Yes	2.7	5.0	5.5	4.7	3.7	5.1
Reproductive factors						
Age at menarche						
<14	38.4	42.3	45.5	39.6	37.1	46.1
15–16	38.6	40.0	38.3	42.7	41.9	37.9
17+	23.1	17.7	16.1	17.8	21.1	16.0
Age at menopause						
Premenopause	55.2	59.9	61.7	63.2	44.9	60.0
<44	7.0	6.6	6.1	10.4	7.1	6.2
45–49	15.9	12.5	11.6	9.4	20.8	11.9
50–54	18.7	17.6	16.5	14.2	23.6	18.8
55+	3.3	3.5	4.1	2.8	3.6	3.1
Age at first birth						
Nullipara	8.9	7.4	7.8	8.2	5.6	7.4
<23	25.0	22.0	19.3	28.0	28.3	25.8
24–30	60.9	63.4	65.0	56.3	60.9	59.7
31+	5.3	7.2	8.0	7.5	5.3	7.1
Period of breast feeding among breastfed women						
1–6	20.3	24.4	24.4	27.8	17.6	30.2
7+	79.7	75.6	75.6	72.2	82.4	69.8
Oral contraceptive						
Never	93.7	92.1	92.3	91.0	88.6	93.1
Ever	6.3	7.9	7.7	9.0	11.4	6.9
Lifestyle factors						
BMI among postmenopausal women						
<25	69.1	68.6	63.8	67.8	75.0	70.4
25–29.9	28.5	27.0	31.0	26.1	22.2	25.7
30+	2.4	4.4	5.2	6.1	2.8	3.9
Exercise						
No	64.8	76.4	78.1	71.1	79.3	76.2
Yes	35.2	23.6	21.9	28.9	20.7	23.8
Alcohol drinking						
No	40.8	65.1	63.8	65.5	64.5	65.2
Yes	59.2	34.9	36.2	34.5	35.5	34.8

BMI = body mass index, HER2 = human epidermal growth factor 2, HR = hormone receptor.  
 \* Includes 689 breast cancer patients with nonclassified types by hormone and HER2 receptors.

The exposure prevalence of breast cancer risk factors nationwide was estimated using the female population database of the Korea National Health and Nutrition Examination Surveys (KNHANES), a nationally representative study in Korea. We used the results from 2005 because a questionnaire regarding female reproductive history was included from 2005. The prevalence of a family history of breast cancer was estimated using the control subjects of the SeBCS<sup>12</sup> because the KNHANES did not measure it. Assuming a lag period from exposure to cancer development, we estimated the risk distribution by standardizing the prevalence in the KNHANES 2005, or the SeBCS to the female population in mid-1990 (Supplementary Table 1, <http://links.lww.com/MD/A776>).

We calculated the PAFs using the modified Levin formula for multiple categories, as proposed by Hanley.<sup>18,19</sup>

$$PAF = \sum p_i(RR_i - 1) / (1 + \sum p_i(RR_i - 1))$$

Although the equation applied prevalence of exposure in a population (p) and relative risks (RRs), we considered the ORs from the case–control study as RRs because breast cancer is a relatively rare disease, and ORs can be used as RR estimates.<sup>20</sup> When we calculated the PAFs, we multiplied the risk factors by 0.32 and 0.68 for postmenopausal women and premenopausal women, respectively, taking into consideration the prevalence of postmenopausal women in Korea estimated using the KNHANES 2005.

**TABLE 2.** Relative Risks (95% Confidence Intervals) of Major Risk Factors for Breast Cancer Subtypes Classified by HR and HER2 in Seoul Breast Cancer Study

	Breast Cancer Subtype According to HR and HER2				
	Total Cancer (N = 3163)*	Luminal A (N = 1363)	Luminal B-HER2 Positive (N = 297)	HER2-overexpres- sion (N = 380)	Triple Negative (N = 434)
Family history of breast cancer					
No	1.00	1.00	1.00	1.00	1.00
Yes	1.55 (1.13–2.12)	1.63 (1.13–2.36)	1.81 (0.96–3.43)	1.18 (0.63–2.22)	1.47 (0.87–2.50)
Reproductive factors					
Age at menarche					
<14	1.94 (1.39–2.71)	2.15 (1.43–3.23)	1.61 (0.75–3.42)	1.98 (1.03–3.80)	2.61 (1.49–4.58)
15–16	1.43 (1.21–1.70)	1.57 (1.27–1.94)	1.43 (0.98–2.09)	1.51 (1.10–2.07)	1.56 (1.13–2.16)
17+	1.00	1.00	1.00	1.00	1.00
Age at menopause among postmenopausal women					
<44	1.00	1.00	1.00	1.00	1.00
45–49	1.02 (0.76–1.37)	1.02 (0.70–1.47)	0.46 (0.25–0.87)	1.48 (0.86–2.55)	1.13 (0.64–1.99)
50–54	1.33 (0.99–1.79)	1.30 (0.90–1.88)	0.68 (0.37–1.25)	1.60 (0.92–2.78)	1.74 (1.00–3.02)
55+	1.39 (1.00–1.95)	1.66 (1.00–2.83)	0.81 (0.30–2.17)	1.37 (0.59–3.17)	1.37 (0.60–3.16)
Age at first birth among parous women					
<23	1.00	1.00	1.00	1.00	1.00
24–30	1.06 (0.90–1.26)	1.29 (1.04–1.60)	0.78 (0.54–1.11)	0.99 (0.73–1.35)	0.76 (0.56–1.02)
31+	1.14 (0.85–1.54)	1.32 (1.00–1.76)	0.96 (0.50–1.84)	0.87 (0.48–1.58)	0.70 (0.42–1.19)
Period of breast feeding among breast-fed women					
1–6	1.43 (1.17–1.74)	1.45 (1.14–1.84)	1.80 (1.21–2.69)	1.10 (0.75–1.63)	2.00 (1.43–2.80)
7+	1.00	1.00	1.00	1.00	1.00
Oral contraceptive					
Never	1.00	1.00	1.00	1.00	1.00
Ever	1.28 (1.01–1.60)	1.26 (1.01–1.71)	1.58 (0.93–2.69)	1.73 (1.12–2.63)	1.06 (0.66–1.70)
Lifestyle factors					
BMI among postmenopausal women					
<25	1.00	1.00	1.00	1.00	1.00
25–29.9	1.02 (0.87–1.20)	1.03 (0.85–1.26)	0.80 (0.55–1.15)	0.68 (0.49–0.93)	1.22 (0.92–1.60)
30+	1.66 (1.04–2.66)	1.96 (1.15–3.33)	1.65 (0.70–3.88)	1.53 (0.69–3.41)	0.96 (0.40–2.32)
Exercise					
No	1.88 (1.61–2.20)	2.05 (1.68–2.52)	1.37 (0.97–1.94)	2.33 (1.66–3.27)	1.80 (1.34–2.43)
Yes	1.00	1.00	1.00	1.00	1.00
Alcohol drinking					
No	1.00	1.00	1.00	1.00	1.00
Yes	1.10 (0.95–1.26)	1.18 (1.01–1.38)	1.09 (0.80–1.49)	1.30 (1.01–1.67)	1.03 (0.79–1.32)

BMI = body mass index, HER2 = human epidermal growth factor 2, HR = hormone receptor.

\*Include 689 breast cancer patients with non-classified types by hormone and HER2 receptors.

## RESULTS

The ORs and 95% confidence intervals (CIs) of each risk factor for total breast cancer and breast cancer subtypes are presented in Table 2. Having a family history of breast cancer (OR = 1.55 [95% CI = 1.13–2.12]), younger age at menarche (OR = 1.94 [95% CI = 1.39–2.71] for ≤14 years old and OR = 1.43 [95% CI = 1.21–1.70] for 15–16 years old), older age at menopause (OR = 1.39 [95% CI = 1.00–1.95] for ≥55 years old), having breastfed ≤6 months (OR = 1.43 [95% CI = 1.17–1.74]), experience of oral contraceptive use (OR = 1.28 [95% CI = 1.01–1.60]), higher BMI after menopause (OR = 1.66 [95% CI = 1.04–2.66] for ≥30 kg/m<sup>2</sup>), and exercise (OR = 1.88 [95% CI = 1.61–2.20]) were significantly associated with increased risk of total breast cancer.

We found heterogeneity in significant risk factors according to breast cancer subtypes. Younger age at menarche and

exercise were associated with 3 subtypes of breast cancer—luminal A, HER2-overexpression, and TN subtypes—and not with luminal B type. Family history of breast cancer, age at first birth, and postmenopausal BMI were associated with only luminal A type. Later age at menopause increased the risk of luminal A and TN breast cancers significantly, but decreased the risk of luminal B type. Having breastfed ≤6 months was associated with luminal A, luminal B, and TN subtypes, and experience of oral contraceptive and alcohol use were significantly associated with luminal A and HER2-overexpression type of breast cancer. Among the risk factors, we found significant heterogeneity in the quantification of risk for exercise (*P* heterogeneity 0.029).

Table 3 shows the PAFs for the present theoretical proportion of breast cancers prevented by removing each risk factor and the different proportions preventable according to the subtype presented as *P* heterogeneity. Regarding total breast

**TABLE 3. Population Attributable Fractions of Major Risk Factors for Breast Cancer Subtypes Classified by HR and HER2 in Seoul Breast Cancer Study**  
**Breast Cancer Subtype According to HR and HER2**

	Total Cancer	Luminal A	Luminal B—HER2 Positive	HER2-overexpression	Triple Negative	P Heterogeneity
<b>Individual factor</b>						
Family history of breast cancer	2.2 (2.1–2.2)	2.5 (2.4–2.5)	3.2 (3.1–3.3)	0.7 (0.7–0.7)	1.8 (1.8–1.8)	0.661
Age at menarche	37.7 (31.0–45.9)	43.1 (33.8–55.0)	30.9 (20.5–46.9)	39.5 (27.1–57.9)	49.4 (34.5–70.9)	0.126
Age at menopause*	3.9 (3.7–4.0)	4.1 (3.9–4.3)	–9.9 (–10.6––9.3)	10.2 (9.3–11.2)	8.2 (7.5–9.0)	<0.001
Age at first birth†	2.8 (2.6–2.9)	10.8 (9.5–12.4)	–10.7 (–13.3– –8.6)	–0.8 (–1.1– –0.5)	–12.8 (–15.3––10.7)	<0.001
Breast feeding duration†	4.2 (4.0–4.3)	4.4 (4.2–4.5)	7.4 (6.8–8.0)	1.0 (1.0–1.0)	9.0 (8.3–9.7)	0.016
Oral contraceptive†	3.3 (3.2–3.4)	3.0 (2.9–3.2)	6.4 (5.7–7.2)	7.9 (7.1–8.8)	0.7 (0.7–0.7)	0.010
Body mass index*	1.0 (1.0–1.0)	1.5 (1.4–1.5)	–0.8 (–0.9––0.8)	–2.0 (–2.1––2.0)	1.7 (1.7–1.8)	0.005
Exercise	28.8 (25.5–32.6)	32.6 (27.7–38.4)	14.5 (11.4–18.6)	38.0 (28.7–50.2)	26.9 (21.4–33.9)	0.014
Alcohol	5.1 (4.8–5.4)	9.2 (8.5–9.9)	4.6 (4.0–5.2)	15.3 (13.4–17.5)	1.5 (1.5–1.6)	<0.001
Modifiable factors, total	45.1 (41.1–49.7)	61.4 (54.3–69.8)	21.4 (18.6–24.9)	59.4 (47.8–74.3)	27.1 (22.9–32.4)	<0.001
<b>Premenopausal women</b>						
Reproductive factors	10.2 (9.8–10.6)	18.2 (16.6–20.0)	3.1 (2.0–3.9)	8.1 (7.7–8.6)	–3.1 (–4.9– –1.6)	<0.001
Lifestyle factors	23.1 (20.6–25.9)	28.4 (24.6–32.8)	13.0 (10.5–16.2)	36.2 (28.6–46.1)	19.3 (15.5–24.2)	0.003
Total	33.3 (30.4–36.5)	46.6 (41.3–52.9)	16.1 (14.5–18.2)	44.4 (36.3–54.7)	16.2 (13.9–19.3)	<0.001
<b>Postmenopausal women</b>						
Lifestyle factors	11.9 (10.7–13.2)	14.8 (13.0–16.9)	5.3 (4.2–6.7)	15.0 (11.5–19.6)	10.8 (9.0–13.1)	0.153
Reproductive factors, total	51.8 (44.5–60.6)	65.4 (54.4–79.3)	24.1 (15.2–38.2)	57.9 (44.0–77.8)	54.5 (40.3–75.0)	<0.001

HER2 = human epidermal growth factor 2, HR = hormone receptor.

\* Considered only for postmenopausal women; In the Korean population, the prevalence of post-menopausal women was 32%.

† Considered only for premenopausal women. In the Korean population, the prevalence of pre-menopausal women was 68%.



cancer, the PAF (95% CI) of the modifiable factors including age at first birth, breast feeding duration, oral contraceptive use, BMI after menopause, exercise, and alcohol drinking was 33.3% (30.4–36.5%) among premenopausal women and 11.9% (10.7–13.2%) among postmenopausal women. Modifiable factors were differently attributed to each subtype (luminal A, PAF = 61.4% [54.3%–69.8%]; luminal B, 21.4% [18.6%–24.9%]; HER2-overexpression, 59.4% [47.8%–74.3%], and TN, 27.1% [22.9%–32.4%],  $P$  heterogeneity  $\leq 0.001$ ). The contribution of modifiable reproductive factors for luminal A type in premenopausal women was higher than that of the other subtypes (18.2% for luminal A; 3.1%, 8.1%, and –3.1% for luminal B, HER2-overexpression, and TN subtypes, respectively;  $P$  heterogeneity  $\leq 0.001$ ), and lifestyle factors were also heterogeneous in premenopausal women (luminal A, 28.3%; luminal B, 13.0%; HER2-overexpression, 36.2%, and TN subtypes, 19.3%;  $P$  heterogeneity = 0.003). However, the PAF for lifestyle factors in postmenopausal women did not show any significant differences ( $P$  heterogeneity = 0.153).

The PAF of each risk factor was also significantly different for each subtype. The PAFs for age at menopause, age at first birth, breastfeeding duration, OC use, BMI after menopause, exercise, and alcohol drinking were significantly different ( $P < 0.05$ ), suggesting significant heterogeneity in the preventable portions of the disease following quantitative risk between subtypes.

When we classified breast cancers according to estrogen and progesterone receptors, the PAFs for age at menopause, age at first birth, BMI after menopause, exercise, and alcohol consumption were significantly different ( $P < 0.05$ ). The PAFs for modifiable factors, including reproductive and lifestyle factors in premenopausal women, were significantly different, ranging from 29.9% to 60.4% (Supplementary Tables 2 and 3, <http://links.lww.com/MD/A776>). However, the differences in the PAFs of the modifiable and reproductive factors were more prominent when we used a subtype classification according to hormone and HER2 receptors.

## DISCUSSION

There were heterogeneities in the quantitative risk followed by the PAFs according to tumor subtypes based on the HR and HER2 status, suggesting differences in theoretically expected reductions in breast cancer from changes in modifiable factors. In the PAF comparisons, premenopausal modifiable factors including reproductive and lifestyle factors were heterogeneous across breast cancer subtypes; whereas non-modifiable factors, such as family history and age at menarche, and modifiable factors in postmenopausal women were not heterogeneous.

Our study population consisted of 55.1%, 12.0%, 15.4%, and 17.5% of luminal A type, luminal B type, HER2 type, and TN type of breast cancer patients, respectively. Previous studies have shown that luminal A type accounted for about half of the breast tumors, and the proportions of the other subtypes in our study were similar with other Asian women.<sup>21,22</sup>

Among total breast cancer cases, 45.1% were attributed to modifiable risk factors including 33.3% for premenopausal women and 11.9% for postmenopausal women. Alteration in exercise habits could potentially reduce female breast cancer incidence in Korea by 29%. A study showed that the PAF of modifiable risk factors was 31.5%, and the most important modifiable factors were HRT use and physical activity (19.4%

and 12.8%) among postmenopausal breast cancers.<sup>5</sup> Another study estimated that “alcohol intake + physical activity” were the most important in premenopausal women, and “physical activity + BMI” were important in postmenopausal women.<sup>23</sup> Sprague et al<sup>24</sup> estimated that summary PAFs were 40.7% for modifiable factors, and weight gain and physical inactivity had the highest PAFs (21.3% and 15.7%), whereas Clarke et al<sup>4</sup> estimated that PAFs were 2% to 11% for HRT use, 1% to 20% for alcohol intake, and 2% to 15% for physical inactivity, and 2% to 11% for breastfeeding. The message that physical activity is the most important modifiable factor as an intervention in breast cancer is consistent in our and previous studies.

According to subtypes, intervention of modifiable factors could reduce the luminal A and HER2 type breast cancers mostly (about 60%) compared with luminal B and TN type breast cancers (about 20%). Exercise might be the most effective lifestyle modification for prevention in all subtypes, but the effects (preventable proportions) varied, ranging from 14.5% to 38.0% according to subtypes, which were also lower in luminal B and TN type breast cancers. Generally, the prognosis for the luminal B and TN subtypes<sup>25,26</sup> was worse, and we found that the preventable proportions of these 2 types through lifestyle modification were lower, especially in premenopausal women. TN type breast cancers are often associated with BRCA1 mutations<sup>27</sup>; thus, it could be suggested that TN type breast cancers may have a strong heritable component. However, considering the prevalence of BRCA1 mutation in TN type breast cancers differed by ethnicity and age,<sup>27</sup> more studies about the smaller contribution of modifiable factors to the development of this type of breast cancer would be necessary in relation with genetic contributions. Although the preventable proportions were lower in these 2 subtypes, it is noteworthy that exercise and breast feeding could prevent 27% and 9% of TN breast cancer cases in which the prognosis is poor, and a standard treatment has not been established yet<sup>28</sup> suggesting that strategies for prevention might be a better way to control TN breast cancers.

The PAFs for reproductive factors in premenopausal women varied. PAFs for a later age at first birth showed an increased contribution only for luminal A cancers. The contribution of breastfeeding duration to HER2-overexpression breast cancer was small, but that of oral contraceptives was greater in HER2 receptor-positive breast cancers, such as luminal B and HER2-overexpression. In particular, the PAF of all reproductive factors for the luminal B type was lower than that of the other subtypes, presenting inverse PAFs for age at menopause, age at first birth, and BMI after menopause. In addition, the PAFs of BMI for postmenopausal HER2 receptor-positive breast cancers, such as luminal B type and HER2-overexpression, were negative, suggesting different risks and contributions for different breast cancer subtypes. These differences in PAFs were affected by various directions and strengths of the associations with risk factors; thus, further studies are needed to identify the biological mechanisms of these inconsistent effects of risk factors according to breast cancer subtype.

Although many previous studies have shown different effects of risk factors on the risk of breast cancer subtypes by receptor status<sup>7,9,10</sup> or morphological type,<sup>29,30</sup> only a few studies have tried to estimate the PAFs according to breast cancer subtypes. One study investigated the quantitative risks of modifiable and nonmodifiable factors and related PAFs according to ER and PR and found significant different risk factors and associated PAFs by receptor status.<sup>5</sup> In addition to ER and PR, HER2 status was considered additionally in the risk assessment

and PAF estimation, and we also found heterogeneity according to subtypes. When we compared our PAF estimates based on ER, PR, and HER2 status and those based on ER and PR status, the differences between subtypes were more distinct when HER2 was also considered, especially for reproductive factors including breastfeeding and OC use (Appendix Table 3, <http://links.lww.com/MD/A776>). Compared with the previous study that suggested modification of lifestyle factors including HRT use, physical activity, BMI, and alcohol drinking might be effective in only the ER+/PR+ type,<sup>5</sup> we suggest that not only hormone receptor status but also HER2 be considered in differentiating breast cancer subtypes. More studies are needed to investigate the ORs and PAFs according to subtypes based on the HRs and HER2 to validate the hypothesis of heterogeneity in etiology according to different breast cancer subtypes.

The estimated PAF for total reproductive factors in our study was similar with previous studies<sup>5,24</sup> but higher than the result from China whose PAF was only 6.7%<sup>31</sup> and lower than that from Iran whose PAF for only parity was >50%.<sup>32</sup> Significantly different PAFs by subtypes identified in this study may explain the differences from previous studies that did not divide subtypes. Among the nonmodifiable reproductive factors, age at menarche was the most important factor for total breast cancer and all subtypes, and it might be caused by the high prevalence rates in the risk categories (About 84% were included in the risk categories, Supplementary Table 1, <http://links.lww.com/MD/A776>).

Among the lifestyle factors, the ORs and PAFs of alcohol consumption and BMI after menopause were quiet inconsistent between studies.<sup>4,5,23,24</sup> These study variations might be caused by not only differences in the prevalence and ORs between populations but also in the cut-points for categorization, which are known to have a large effect on PAF.<sup>33</sup>

This study has several limitations. We applied prevalence rates in 2005 using a nationwide study by standardizing the population in 1990. Considering the rapid changes in reproductive factors and westernized lifestyle, our results might underestimate the PAF for modifiable and nonmodifiable factors in Korea.<sup>34</sup> We calculated the associated factors with breast cancer and their risk using a case-control study, and there would be selection bias or recall bias, or limited representativeness. These biases would cause nondifferential exposure misclassification, and the PAFs would be underestimated.<sup>35</sup> When we classified breast cancer according to subtypes, we could not consider Ki67 because of limited information. Although the PAFs of individual factors cannot be summed for risk factor combinations and cannot be subtracted from 100% to determine unexplained proportions of diseases,<sup>36</sup> we added the PAF of each modifiable and nonmodifiable factor to estimate the proportion of breast cancers preventable through the elimination of modifiable factors intuitively. The study recruitment period was 7 years; thus, the reproductive and lifestyle factors may have changed. However, we did not conduct a stratified analysis according to the time interval because the numbers of patients with luminal B-HER2 positive, HER2-over expression, and triple-negative cancers were relatively small, and a stratified analysis might not be appropriate. Although risk factors differ by breast cancer subtype, we applied the same model to all subtypes, as described by Barnes et al<sup>5</sup> because we were interested in determining whether the effects of known risk factors are similar across cancer subtypes. In addition, we matched all breast cancer cases and controls, but we used all of the controls for each subtype and did not consider matching when conducting the analysis according to subtypes

classified by HR and HER2 receptors. To the best of our knowledge, this is the first study to estimate the risks and PAFs by hormone receptors and HER2 including various risk factors, suggesting a heterogeneity regarding tumor subtypes.

In conclusion, a substantial proportion of breast cancers could be reduced by eliminating modifiable risk factors. However, breast cancer is a heterogeneous disease, and HR and HER2 status may be associated with heterogeneous risk factors and their attributable risk, suggesting a different etiology. An intervention for modifiable factors could be effective for breast cancer prevention, especially in premenopausal women. Additionally, exercise is the most effective strategy, although luminal B and TN subtypes seemed to be less preventable.

## ACKNOWLEDGMENTS

The authors thank Mathieu Boniol in International Prevention Research Institute and Paolo Boffetta in Western Pacific Regional Office, World Health Organization for helping us to start this project.

## REFERENCES

1. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127:2893–2917.
2. Key TJ, Verkasalo PK, Banks E. Epidemiology of breast cancer. *Lancet Oncol*. 2001;2:133–140.
3. Northridge ME. Public health methods—attributable risk as a link between causality and public health action. *Am J Public Health*. 1995;85:1202–1204.
4. Clarke CA, Purdie DM, Glaser SL. Population attributable risk of breast cancer in white women associated with immediately modifiable risk factors. *BMC Cancer*. 2006;6:170.
5. Barnes BB, Steindorf K, Hein R, et al. Population attributable risk of invasive postmenopausal breast cancer and breast cancer subtypes for modifiable and non-modifiable risk factors. *Cancer Epidemiol*. 2011;35:345–352.
6. Granstrom C, Sundquist J, Hemminki K. Population attributable risks for breast cancer in Swedish women by morphological type. *Breast Cancer Res Treat*. 2008;111:559–568.
7. Althuis MD, Fergenbaum JH, Garcia-Closas M, et al. Etiology of hormone receptor-defined breast cancer: a systematic review of the literature. *Cancer Epidemiol Biomarkers Prev*. 2004;13:1558–1568.
8. Parise CA, Caggiano V. Breast cancer survival defined by the ER/PR/HER2 subtypes and a surrogate classification according to tumor grade and immunohistochemical biomarkers. *J Cancer Epidemiol*. 2014;2014:469251.
9. Ma H, Bernstein L, Pike MC, et al. Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies. *Breast Cancer Res*. 2006;8:R43.
10. Yang XR, Chang-Claude J, Goode EL, et al. Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies. *J Natl Cancer Inst*. 2011;103:250–263.
11. Aebi S, Davidson T, Gruber G, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2011;22(Suppl 6):vi12–vi24.
12. Park B, Ma S, Shin A, et al. Korean risk assessment model for breast cancer risk prediction. *PLoS One*. 2013;8:e76736.
13. Lee JS, Oh M. Reproductive factors and subtypes of breast cancer defined by estrogen receptor, progesterone receptor, and human

- epidermal growth factor receptor 2: a register-based study from Korea. *Clin Breast Cancer*. 2014;14:426–434.
14. Grosse Y, Baan R, Straif K, et al. A review of human carcinogens—Part A: pharmaceuticals. *Lancet Oncol*. 2009;10:13–14.
  15. Secretan B, Straif K, Baan R, et al. A review of human carcinogens—Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. *Lancet Oncol*. 2009;10:1033–1034.
  16. Koh K. Estimates of alcohol consumption based on pure alcohol content and their comparison with data of other countries. *Health and Welfare Policy Forum*. 2000;6:77–86.
  17. Kim HC, Oh SM. Noncommunicable diseases: current status of major modifiable risk factors in Korea. *J Prevent Med Public Health*. 2013;46:165–172.
  18. Hanley JA. A heuristic approach to the formulas for population attributable fraction. *J Epidemiol Commun Health*. 2001; 55:508–514.
  19. Levin ML. The occurrence of lung cancer in man. *Acta Unio Int Contra Cancrum*. 1953;9:531–541.
  20. Schmidt CO, Kohlmann T. When to use the odds ratio or the relative risk? *Int J Public Health*. 2008;53:165–167.
  21. Clarke CA, Keegan TH, Yang J, et al. Age-specific incidence of breast cancer subtypes: understanding the black-white crossover. *J Natl Cancer Inst*. 2012;104:1094–1101.
  22. Iwase H, Kurebayashi J, Tsuda H, et al. Clinicopathological analyses of triple negative breast cancer using surveillance data from the Registration Committee of the Japanese Breast Cancer Society. *Breast cancer (Tokyo, Japan)*. 2010;17:118–124.
  23. Mezzetti M, La Vecchia C, Decarli A, et al. Population attributable risk for breast cancer: diet, nutrition, and physical exercise. *J Natl Cancer Inst*. 1998;90:389–394.
  24. Sprague BL, Trentham-Dietz A, Egan KM, et al. Proportion of invasive breast cancer attributable to risk factors modifiable after menopause. *Am J Epidemiol*. 2008;168:404–411.
  25. Ades F, Zardavas D, Bozovic-Spasojevic I, et al. Luminal B breast cancer: molecular characterization, clinical management, and future perspectives. *J Clin Oncol*. 2014;32:2794–2803.
  26. Cogliano V, Grosse Y, Baan R, et al. Carcinogenicity of combined oestrogen-progestagen contraceptives and menopausal treatment. *Lancet Oncol*. 2005;6:552–553.
  27. Greenup R1, Buchanan A, Lorizio W, et al. Prevalence of BRCA mutations among women with triple-negative breast cancer (TNBC) in a genetic counseling cohort. *Ann Surg Oncol*. 2013;20:3254–3258.
  28. Zhao S, Ma W, Zhang M, et al. High expression of CD147 and MMP-9 is correlated with poor prognosis of triple-negative breast cancer (TNBC) patients. *Med Oncol (Northwood, London, England)*. 2013;30:335.
  29. Li CI, Daling JR, Malone KE, et al. Relationship between established breast cancer risk factors and risk of seven different histologic types of invasive breast cancer. *Cancer Epidemiol Biomarkers Prevent*. 2006;15:946–954.
  30. Rosenberg LU, Magnusson C, Lindstrom E, et al. Menopausal hormone therapy and other breast cancer risk factors in relation to the risk of different histological subtypes of breast cancer: a case-control study. *Breast Cancer Res*. 2006;8:R11.
  31. Li L, Ji J, Wang JB, et al. Attributable causes of breast cancer and ovarian cancer in china: reproductive factors, oral contraceptives and hormone replacement therapy. *Chin J Cancer Res*. 2012;24:9–17.
  32. Ghiasvand R, Bahmanyar S, Zendehdel K, et al. Postmenopausal breast cancer in Iran; risk factors and their population attributable fractions. *BMC Cancer*. 2012;12:414.
  33. Rockhill B, Weinberg CR, Newman B. Population attributable fraction estimation for established breast cancer risk factors: considering the issues of high prevalence and unmodifiability. *Am J Epidemiol*. 1998;147:826–833.
  34. Yoo KY, Kim Y, Park SK, et al. Lifestyle, genetic susceptibility and future trends of breast cancer in Korea. *Asian Pac J Cancer Prev*. 2006;7:679–682.
  35. Hsieh CC, Walter SD. The effect of non-differential exposure misclassification on estimates of the attributable and prevented fraction. *Stat Med*. 1988;7:1073–1085.
  36. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *Am J Public Health*. 1998;88:15–19.