

Estrogen-DNA adduct ratios as a predictor for breast cancer risk in premenopausal Asian women

Ching-Hung Lin

National Taiwan University Hospital, Cancer Center Branch

Muhammad Zahid

University of Nebraska Medical Center

Wen-Hung Kuo (✉ wenhung003915@ntu.edu.tw)

National Taiwan University Hospital

Fu-Chang Hu

National Taiwan University

Ming-Yang Wang

National Taiwan University Hospital

I-Chun Chen

National Taiwan University Hospital, Cancer Center Branch

Cheryl L. Beseler

University of Nebraska Medical Center

Bodhisattwa Mondal

University of Nebraska Medical Center

Yen-Shen Lu

National Taiwan University Hospital

Eleanor G. Rogan

University of Nebraska Medical Center

Ann-Lii Cheng

National Taiwan University Hospital

Research Article

Keywords: breast cancer, estrogen, estrogen-DNA adduct, premenopausal

Posted Date: June 16th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1749357/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

The incidence of breast cancer among East Asian women has been increasing rapidly in recent decades, but the etiology remains unclear. This case-control study investigated whether estrogen-DNA adducts ratio is a predictor of breast cancer risk in East Asian women.

Methods

The control group ($n = 146$) comprised healthy female volunteers and women with benign breast disease without atypia. The case group ($n = 221$) comprised women either with benign breast disease with atypia and breast cancer. The ratios of estrogen-DNA adducts to their respective metabolites and conjugates in plasma were analyzed using ultraperformance liquid chromatography and tandem mass spectrometry. The single nucleotide polymorphisms (SNPs) of *CYP1A1*, *CYP1B1*, and *COMT* were genotyped, and a logistic regression model was used to compare the estrogen-DNA adduct ratios between the two groups.

Results

The estrogen-DNA adduct ratio in the case group was significantly higher than that in the control group (median ratio: 58.52 vs. 29.36, $P = 0.004$). A multiple logistic regression model demonstrated that a unit increase in the natural log of the estrogen-DNA adduct ratio in premenopausal women was a significant predictor of breast cancer risk, with an estimated hazard ratio of 1.718 (1.444 – 2.046, $P < 0.001$). However, the *CYP1A1*, *CYP1B1*, and *COMT* SNPs were not associated with the estrogen-DNA adduct ratios.

Conclusion

Plasma estrogen-DNA adduct ratio predicts breast cancer risk in premenopausal women in East Asia, and the genotoxic effects may contribute to the carcinogenesis of this emerging disease.

Background

The incidence of breast cancer among women in Asian countries is generally lower than that in Western countries.[1] However, the incidence has increased rapidly in East Asia over the past 40 years, including in Singapore, Taiwan, Korea, Japan, and China [2]. Previous age-period-cohort analyses have demonstrated a strong cohort effect on breast cancer incidence in these countries, with an increase observed particularly among women aged < 50 years [3–7], and our studies have identified an unusually high prevalence of hormone receptor-positive breast tumors in patients aged ≤ 50 years in Taiwan and other East Asian countries [2, 8, 9].

Notably, in addition to hormone receptor-positive breast cancer, our previous study observed that the incidences of type I endometrial cancer and endometrioid carcinoma of the ovary have increased rapidly in Taiwan in recent decades [10]. The common features of these cancers include high rates of hormone receptor expressions and their occurrences are considered to be associated with estrogen exposure. By contrast, the incidence of other subtypes of these three cancers and other major cancers, such as lung and colon cancer, have also increased, but mainly in older adults [10]. These findings indicate that estrogen-related malignancies are emerging in young women in East Asia, and estrogen-related etiologies, either exogenous estrogenic substances or endogenous estrogen, are linked to the rapid increase in incidence.

With regard to endogenous estrogen, early studies (published between 1971 and 1991) of premenopausal women found that estradiol levels were 20–50% lower in Asian women (including Asian American women) than in Caucasian women of comparable age [11–15]. However, a recent study demonstrated no difference in serum estradiol levels between premenopausal Asian American women and Caucasian women [16], and another recent study even identified higher estradiol levels in premenopausal Asian American women than in Caucasian women [17]. These indirect pieces of evidence suggest that serum estrogen levels have increased considerably in premenopausal Asian American women in the past two decades.

In the present study, we hypothesized that the rapid increase in estrogen levels in premenopausal women in East Asia has led to the production of mutagenic metabolites. Estrogens are oxidized to form catechol estrogens, 2-hydroxyestrone (estradiol) [2-OHE1(E2)], and 4-OHE1(E2) in the presence of cytochrome P450 (CYP) 1A1 and 1B1. These catechol estrogens are further converted to semiquinones and quinones as a result of peroxidase activity. Catechol estrogen-3,4-quinones react with DNA to form depurinating estrogen-DNA adducts 4-OHE1(E2)-1-N3Ade and 4-OHE1(E2)-1-N7Gua. The apurinic sites of the reacted DNA can lead to mutations caused by error-prone DNA repair, initiating cancer [18]. Depurinating estrogen-DNA adducts are shed into the blood and excreted in urine. Previous studies have demonstrated that in Western countries, the ratio of estrogen-DNA adducts to estrogen metabolites and conjugates in blood or urine is significantly higher in women at high risk of breast cancer or who have been diagnosed with the disease than in women at average risk of breast cancer [19, 20].

This study investigated the association between estrogen-DNA adduct ratio and breast cancer risk in Taiwanese women, and the analysis was stratified by menopausal status to uncover this association among premenopausal women. In addition, we evaluated the interaction of single nucleotide polymorphisms (SNPs) of *CYP1A1*, *CYP1B1*, and *catechol-O-methyltransferase (COMT)* with estrogen-DNA adduct ratio and the interaction effect on breast cancer risk. *COMT* is a phase II enzyme and can mitigate the harmful effects of catechol estrogens by catalyzing the O-methylation in catechol estrogens to methyl estrogens [21].

Methods

Study population

In October 2010, we initiated a case-control study to evaluate the associations of 1) estrogenic environmental pollutants and 2) estrogen synthesis and metabolism with breast cancer risk in Taiwanese women. The study was approved by the Ethics Committee of National Taiwan University Hospital (NTUH; 201004060R). The participants included female volunteers with no history of malignancy or breast disease (defined as healthy volunteers), biopsy-proven benign breast disease, or breast cancer patients. Healthy volunteers were recruited through posters and flyers at NTUH and in the community, and they received a small fee for their time (approximately US \$6.30) after completing the study. The diagnosis of breast cancer (including ductal carcinoma in situ) and benign breast disease was confirmed by histological samples. Benign breast diseases were classified into proliferative and non-proliferative diseases. Proliferative disease was indicated if specimens contained any of the following: atypical ductal hyperplasia, atypical lobular hyperplasia, ductal hyperplasia (greater than mild), papilloma, radial scars, or sclerosing adenosis. Cysts, fibroadenoma, or columnar changes were considered non-proliferative disease [22, 23]. A meta-analysis demonstrated that proliferative disease was associated with breast cancer risk [24] but that non-proliferative disease was not; thus, we included female volunteers with no history of malignancy and women with non-proliferative breast disease into a control group, and we included women with proliferative breast disease and breast cancer into a case group.

At enrollment, the participants provided a fasting blood sample and completed an assisted questionnaire. To reduce treatment-related bias, blood samples were collected before any treatment for breast cancer was initiated. Blood samples were centrifuged immediately after collection, and the plasma samples were stored at -80°C . The questionnaire included questions on the risk factors for breast cancer, including age, parity, breastfeeding history, body mass index (BMI), smoking and drinking habits, menopause status, education level, and family history of breast or ovarian cancer.

Sample preparation and ultraperformance liquid chromatography and tandem mass spectrometry analysis of estrogen-DNA adduct ratio

Serum samples were stored at -80°C and thawed only once before analysis. Serum aliquots (0.5 ml) were partially purified by solid-phase extraction (SPE) with a small modification, as previously described [19]. The serum samples were diluted with an equal volume of 0.01 M ammonium formate buffer (pH 7; loading and washing buffer) before being passed through the SPE cartridges. The cartridges were washed with methanol and distilled water and, finally, preconditioned with loading buffer. The samples were loaded and passed without vacuum, and the target compounds were eluted by using elution buffer (methanol/acetonitrile/water, 80:10:10 pH 3.5). The eluted samples were lyophilized and reconstituted in 50 μl of methanol/water (0.1% formic acid), and passed through 5000 cutoff filters.

The ultraperformance liquid chromatography and tandem mass spectrometry analyses were carried out using a Waters Acquity UPLC system connected to a high-performance Quattro Micro triple quadrupole mass spectrometer (Waters, Milford, MA, USA) [25]. The analysis was performed by injecting 10 μl of each partially purified serum sample into an Acquity UPLC BEH C18 1.7-mm column ($1 \times 100 \text{ mm}^2$) at a flow rate of 0.15 mL/min. Analytes were identified by their retention time and fragmentation pattern and processed using QuanLynx software to quantify the estrogen-DNA adducts, estrogen metabolites, and conjugates. The adduct ratio was defined by the following equation:

$$= \left[\frac{4 - \text{OHE1 (E2)} - 1 - \text{N3Ade} + 4 - \text{OHE1 (E2)} - 1 - \text{N7Gua}}{4 - \text{catecholestrogens} + 4 - \text{catecholestrogenconjugates}} \right] + \frac{2 - \text{OHE1 (E2)} - 1 - \text{N3Ade}}{4 - \text{catecholestrogens} + 4 - \text{catecholes}}$$

The estrogen metabolites included 2-OHE1(E2) and 4-OHE1(E2). The conjugates included 2-OCH3E1(E2), 4-OCH3E1(E2), 4-OHE1(E2)-2-SG, 4-OHE1(E2)-2-Cys, and 4-OHE1(E2)-2-NAcCys. The adducts included 4-OHE1(E2)-1-N3Ade, 4-OHE1(E2)-1-N7Gua, and 2-OHE1(E2)-6-N3Ade. The concentration of each of the 20 compounds and the ratio of depurinating N3Ade and N7Gua adducts to the sum of their respective estrogen metabolites and conjugates in each serum sample provided a ratio that reflected the degree of imbalance in that estrogen metabolism that could lead to the initiation of cancer. The experiment was a blinded one.

Genotyping of CYP1A1, CYP1B1, and COMT

The genomic DNAs of the blood and tumor specimens were extracted using a QIAamp DNA Mini Kit (Qiagen, Valencia, CA, USA). Genotypes for the *CYP1A1* Ile462Val (rs1048943), *CYP1B1* Leu432Val (rs1056836), and *COMT* Val158Met (rs4680) polymorphisms were determined by a restriction fragment-length polymorphism assay. The polymerase chain reaction (PCR) primers had been designed previously: *CYP1A1* (rs1048943, A2455G): forward 5'-CTG TCT CCC TCT GGT TAC AGG AAGC - 3' and reverse 5'-TTC CAC CCG TTG CAG CAG GAT AGCC-3'; *CYP1B1* (rs1056836, C1294G): forward 5'-CAC TGC CAA CAC CTC TGT CT-3' and reverse 5'-GCA GGC TCA TTT GGG TTG-3'; and *COMT* (rs4680, G472A): forward 5'-CGA GGC TCA TCA CCA TCG AGA TC-3' and reverse 5'-CTG ACA ACG GGT CAG GAA TGC A-3'. These PCR primers were used to generate 204 bp (*CYP1A1*), 294 bp (*CYP1B1*), and 108 bp (*COMT*) PCR products containing SNP sites [26–28]. The PCR products were digested with restriction enzymes (FastDigest BseMI/*CYP1A1* rs1048943, FastDigest Eco57I/*CYP1B1* rs1056836, and FastDigest Hin1II/*COMT* rs4680) according to the manufacturer's protocol (Thermo Fisher Scientific, USA). The three different SNP genotypes of each gene were determined by DNA fragment length on 3% agarose gel as follows: *CYP1A1* (rs1048943, A2455G): A/A homozygotes (149 and 55 bp), A/G heterozygotes (204, 149, and 55 bp), and G/G homozygotes (204 bp); *CYP1B1* (rs1056836, C1294G): C/C homozygotes (187 and 107 bp), C/G heterozygotes (294, 187, and 107 bp), and G/G homozygotes (294 bp); and *COMT* (rs4680, G472A): G/G homozygotes (108 bp), G/A heterozygotes (108, 72, and 36 bp), and A/A homozygotes (72 bp and 36 bp).

Statistical analysis

The categorical variables were described in terms of the frequency and percentage, and the continuous variables were described in terms of the mean \pm standard deviation and median. Power transformations (x^q), including the natural logarithm ($q = 0$), square root ($q = 0.5$), and square ($q = 2$), were applied to some continuous variables for their distributions to be more symmetrical.

A univariate analysis was performed to examine the differences in the distributions of the continuous and categorical variables between the case and control groups using a two-sample Student t test, Wilcoxon rank-sum test (or Mann–Whitney U test), chi-square test, and Fisher exact test (if the expected values in any of the cells of the contingency table was < 5), as appropriate for the data type. A multivariate analysis was then conducted using a fitted linear regression model and logistic regression model to estimate the adjusted effects of potential risk factor predictors on continuous or binary outcomes. All statistical analyses were performed in R software (version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was indicated by a two-sided P value ≤ 0.05 . The statistical analysis is detailed in the Supplementary Materials and Data.

Results

Demographics of the study population

The control group included 88 healthy volunteers and 58 women with non-proliferative benign breast disease. The case group included 33 women with proliferative benign breast disease and 188 patients with breast cancer ($n = 150$) or ductal carcinoma in situ ($n = 38$), diagnosed at NTUH. Participants were recruited between February 2010 and November 2014. Table 1 presents the demographic characteristics of the case and control group participants. The case group had lower percentages of participants at menopause (control: 46%; case: 39%, $P = 0.027$) and a nonsignificantly higher BMI (BMI, 23.2 vs. 22.5 kg/m², $P = 0.098$) than the control group. The median age, smoking habits, alcohol consumption, pregnancy history, number of births, breastfeeding history, education level, and family history of breast cancer were not significantly different between the two groups.

Table 1
Demographic characteristics of control (healthy volunteers and benign low risk) and case (breast cancer and benign high risk)

	Control group			Case group			P*
	Healthy volunteers (n = 88)	Non-proliferative disease (n = 58)	Whole-controls (n = 146)	Proliferative disease (n = 33)	Breast cancer or carcinoma in situ (n = 188)	Whole-cases (n = 221)	
Age (median, range)	48 (27–69)	48 (26–67)	48 (26–69)	45 (27–63)	49 (28–77)	48 (27–77)	0.909
Menopause (n, %)	42 (48%)	25 (43%)	67 (46%)	7 (21%)	69 (37%)	76 (39%)	0.027
Body mass index (kg/m ²) (mean, SD)	22.4 ± 3.4	22.5 (3.1)	22.5 (3.3)	22.6 (3.6)	23.3 (3.8)	23.2 (3.7)	0.098
Cigarette smoking (n, %)	0 (0%)	4 (7%)	4 (3%)	4 (12%)	10 (5%)	14 (6%)	0.143
Alcohol drinking (n, %)	4 (5%)	4 (7%)	8 (6%)	1 (3%)	16 (9%)	17 (8%)	0.527
Pregnant history (n, %)	62 (70%)	49 (84%)	111 (76%)	21 (64%)	149 (79%)	170 (77%)	0.469
No. of birth							0.818
0	28 (32%)	10 (17%)	38 (26%)	14 (42%)	46 (24%)	60 (27%)	
1	11 (13%)	9 (16%)	20 (14%)	5 (15%)	31 (16%)	36 (16%)	
2	29 (33%)	27 (47%)	56 (38%)	9 (27%)	75 (40%)	84 (38%)	
≥ 3	20 (23%)	12 (21%)	32 (22%)	5 (15%)	36 (19%)	41 (19%)	
Breastfeeding (n, %)	39 (44%)	29 (50%)	68 (47%)	10 (30%)	85 (45%)	95 (43%)	0.498
Education level ≥ 12 years (n, %)	88 (100%)	51 (88%)	139 (95%)	31 (94%)	176 (94%)	207 (94%)	0.534
Family history of breast or ovarian cancer (n, %)	9 (10%)	6 (10%)	15 (10%)	7 (21%)	25 (13%)	32 (14%)	0.238
<i>CYP1A1</i> (rs1048943) (n, %)							
Ile/Ile	45 (51%)	30 (52%)	75 (51%)	20 (61%)	97 (52%)	117 (53%)	0.571
Ile/Val	36 (41%)	23 (40%)	59 (40%)	11 (33%)	81 (43%)	92 (42%)	
Val/Val	7 (8%)	5 (9%)	12 (8%)	2 (6%)	10 (5%)	12 (5%)	
<i>CYP1B1</i> (rs1056836)							
leu/leu	67 (76%)	46 (79%)	113 (77%)	25 (76%)	152 (81%)	177 (80%)	0.249
leu/Val	21 (24%)	12 (21%)	33 (23%)	8 (24%)	33 (18%)	41 (19%)	
Val/Val	0 (0%)	0 (0%)	0 (0%)	0 (0%)	3 (2%)	3 (1%)	
<i>COMT</i> (rs4680)							
Val/Val	44 (50%)	35 (60%)	79 (54%)	14 (42%)	102 (54%)	116 (52%)	0.825
Val/Met	32 (36%)	22 (38%)	54 (37%)	15 (45%)	73 (39%)	88 (40%)	
Met/Met	12 (14%)	1 (2%)	13 (9%)	4 (12%)	13 (7%)	17 (8%)	

* The sample statistics presented in this table were mean ± standard deviation (SD) for continuous variables and frequency (percentage, %) for categorical variables. The listed *p*-values of statistical tests were calculated using the Wilcoxon rank-sum test for continuous variables and the Pearson's Chi-squared test for categorical variables..

Correlations of estrogen-DNA adduct ratios with breast disease status and demographic variables

The ratios of estrogen-DNA adduct to these selected categorical variables are listed in Table 2. The estrogen-DNA adduct ratios were significantly higher in the case group than in the control group (mean ratios, 58.52 vs. 29.36, *P* = 0.004 using the Student *t* test and *P* < 0.0001 using the Mann–Whitney *U* test, Fig. 1A), but no significant difference in ratio was demonstrated between the healthy volunteers and the women with non-proliferative breast disease and between the women with breast cancer and those with proliferative breast disease (data not shown).

Table 2
The associations of estrogen-DNA adduct ratio with selected variables

	Estrogen-DNA adduct ratio (mean ± SD)	p value*	p value [†]
Study group		0.004	< 0.0001
Control group (n = 146)	29.36 ± 4.19		
Case group (n = 221)	58.52 ± 7.67		
Age group (years)		0.006	0.013
< 40 (n = 112)	72.14 ± 10.98		
40–49 (n = 97)	32.17 ± 4.87		
50–59 (n = 101)	43.44 ± 11.92		
≥ 60 (n = 157)	28.67 ± 3.37		
Menopausal status		0.275	0.417
Premenopausal (n = 224)	51.25 ± 6.04		
Postmenopausal (n = 143)	40.14 ± 8.52		
Alcohol consumption			
No (n = 342)	44.27 ± 4.24	0.048	0.427
Yes (n = 25)	83.17 ± 44.17		
No. of birth		0.019	0.467
0 (n = 98)	60.47 ± 13.89		
1 (n = 56)	58.91 ± 15.81		
2 (n = 140)	41.56 ± 5.25		
≥ 3 (n = 73)	29.83 ± 4.64		
<i>CYP1A1</i> (rs1048943)		0.692	0.676
Ile/Ile (n = 192)	44.19 ± 5.68		
Ile/Val (n = 151)	52.13 ± 9.44		
Val/Val (n = 24)	36.02 ± 12.87		
<i>CYP1B1</i> (rs1056836)		0.690	0.296
Ile/Ile (n = 290)	46.67 ± 5.38		
Ile/Val (n = 74)	47.01 ± 12.67		
Val/Val (n = 3)	68.73 ± 46.21		
<i>COMT</i> (rs4680)		0.694	0.025
Val/Val (n = 194)	41.28 ± 4.97		
Val/Met (n = 142)	54.32 ± 9.99		
Met/Met (n = 29)	48.61 ± 20.32		
Breast cancer subgroup		0.578	0.645
ER+/ HER2- (n = 99)	56.22 ± 13.13		
ER+/ HER2+ (n = 25)	63.57 ± 19.91		
ER-/ HER2- (n = 7)	66.72 ± 34.97		
ER-/ HER2+ (n = 19)	103.22 ± 43.25		
*Statistical significance by two-tailed Students' t-test or ANOVA test			
[†] The sample statistics presented in this table were mean ± standard deviation (SD) for continuous variables and frequency (percentage, %) for categorical variables. The listed p-values of statistical tests were calculated using the Wilcoxon rank-sum test.			

The estrogen-DNA adduct ratio varied between age groups (heterogeneity among age groups: $P = 0.006$ using an analysis of variance [ANOVA] and $P = 0.013$ using the Kruskal–Wallis test), and the group aged < 40 years had the highest ratio (mean ratios: age < 40 years: 72.14; age 40–49 years: 32.17; age 50–59 years: 43.44; age ≥ 60 years: 28.67). Alcohol consumption was significantly associated with higher estrogen-DNA adduct ratios (mean ratios: 83.17 vs. 44.27)

in the Student t test ($P = 0.048$) but not in the Mann–Whitney U test ($P = 0.427$). A lower number of births was significantly associated with increased estrogen-DNA adduct ratios (mean ratios: no parity: 60.47; one birth: 58.91; two births: 41.56; three births or more: 29.83) in the ANOVA (heterogeneity between age groups: $P = 0.019$), but the association was not significant in the Kruskal–Wallis test ($P = 0.467$).

Pearson and Spearman rank correlation coefficients were used to analyze the correlation of estrogen-DNA adduct ratios with continuous variables. The Pearson correlation coefficient revealed that the decreases of age or number of birth was associated with increase of estrogen-DNA adduct ratio, but the Spearman rank correlation coefficient did not reveal any significant association of these variables with estrogen-DNA adduct ratio (Supplementary Table 1).

Multivariate analysis of predictors for breast cancer or proliferative breast disease

A multivariate analysis demonstrated that premenopausal women and per a natural log unit increase in the estrogen-DNA adduct ratio (estimated OR = 1.718, $P < 0.001$), premenopausal women and women aged 39–52 years (estimated OR = 0.293, $P < 0.001$), postmenopausal women and women aged ≥ 61 years (estimated OR = 3.047, $P = 0.002$), and per unit increase in BMI (estimated OR = 1.073, $P = 0.040$) were associated with the risk of breast cancer or proliferative breast disease.

Considering the interaction of menopausal status with estrogen-DNA adduct ratios in predicting disease, we performed a stratified analysis to compare the estrogen-DNA adduct ratios between the case and control groups by menopausal status. Among the premenopausal participants, the estrogen-DNA adduct ratio was significantly higher in the case group than in the control group (mean ratios: 65.65 vs. 24.82, $P = 0.001$ with the Student t test and $P < 0.0001$ using the Mann–Whitney U test; Fig. 1B). In the postmenopausal participants, the estrogen-DNA adduct ratios were significantly higher in the case group than in the control group (mean ratios: 44.92 vs. 34.72, $P = 0.596$ with the Student t test and $P = 0.285$ using the Mann–Whitney U test; Fig. 1C). The estrogen-DNA adduct ratios were not significantly different between healthy volunteers and participants with non-proliferative benign breast disease in the control group, and between participants with proliferative benign breast disease and breast cancer in the case group (data not shown).

Correlations of estrogen-DNA adduct ratios with estrogen metabolism-related SNPs or breast cancer pathological features

The distributions of the *CYP1A1*, *CYP1B1*, and *COMT* SNPs were not significantly different between the case and control groups (Table 1). The *CYP1A1* and *CYP1B1* SNPs were not significantly associated with estrogen-DNA adduct ratios. The estrogen-DNA adduct ratios varied between the three *COMT* SNPs (rs4680) (heterogeneity among SNP groups: $P = 0.025$ using the Kruskal–Wallis test), and the Val/Met group had the highest ratio (mean ratios: Val/Val: 41.28, Val/Met: 54.32, Met/Met: 48.61). Among the breast cancer patients, the estrogen-DNA adduct ratios were not associated with the distributions of positivity of estrogen receptor (ER) or human epidermal growth factor receptor 2 (HER2) in tumor cells. (Table 2).

Discussion

This case–control study demonstrated that plasma estrogen-DNA adduct ratios were associated with breast cancer in Taiwan. The multiple logistic regression model indicated that a unit increase in the natural log of the estrogen-DNA adduct ratio was a significant predictor of breast cancer risk in premenopausal women, with an estimated hazard ratio of 1.718. Age, alcohol consumption, and the number of births were associated with estrogen-DNA adduct ratios with both a Student t test and ANOVA, but only age was statistically significant with the Mann–Whitney U test or Kruskal–Wallis test. Of the three estrogen metabolism-related genes, only the *COMT* SNP was associated with estrogen-DNA adduct ratios using the Kruskal–Wallis test. With regard to breast cancer, the ER/HER2 subtypes were not associated with estrogen-DNA adduct ratio.

In postmenopausal women, studies have demonstrated a positive association between breast cancer risk and circulating concentrations of estrogens. A reanalysis of nine prospective studies indicated that the relative risk of breast cancer for postmenopausal women with increasing quintiles of estradiol concentrations, relative to the lowest quintile, were 1.42, 1.21, 1.80, and 2.00 ($P_{\text{trend}} < 0.001$) [29]. For premenopausal women, evidence is limited because hormone measurements are complicated by the dynamic change of serum estrogen levels across menstrual cycles. However, a collaborative reanalysis of seven prospective studies demonstrated that breast cancer risk was associated with a doubling in concentrations of estradiol (OR 1.19, 95% CI: 1.06–1.35), but the association in premenopausal women was modest, and these seven studies were conducted in Western countries [30]. The present study indicated a strong association between estrogen-DNA adduct ratios and breast cancer risk in premenopausal women. In contrast to serum estrogen levels, the timing of the blood sampling for estrogen-DNA adduct measurements was not limited by menstruation cycles

A pilot study published in 2008 demonstrated that urinary estrogen-DNA adduct ratios were higher in women at high risk of breast cancer ($n = 12$) and with breast cancer ($n = 17$) than in control group participants ($n = 46$). [20] This finding was validated in a study with a larger number of participants (women at high risk of breast cancer: $n = 40$; women with breast cancer: $n = 40$; controls: $n = 40$) [31]. Serum samples were measured in a third study, demonstrating that an estrogen-DNA adduct ratio ≥ 77 predicted breast cancer risk in logistic regression models (women at high risk of breast cancer: $n = 63$; women with breast cancer: $n = 80$; controls: $n = 79$) [19].

The present study differed from the previous three studies with respect to participants, sample type, and the subgroup analysis by menopausal status. With regard to participants, the previous three studies enrolled women from Italy and the United States whereas the present study enrolled women from Taiwan. The differences in estrogen profiles and metabolism between Asian, Asian American, and Caucasian women have been demonstrated in previous studies [32–35]; therefore, conducting a large-scale case–control study with Asian participants is crucial to verify whether estrogen-DNA adduct ratio can serve as a predictive biomarker for breast cancer risk in Asian women. In addition, a 5-year Gail Model score of $\geq 1.66\%$ was used to define a high breast cancer risk in the previous three studies, but in the present study, histology was used to identify proliferative disease, which is a more stringent criteria for defining high risk groups. Regarding sample type, plasma samples were used in the present study whereas urine samples or serum samples were used in the previous three studies. Furthermore, we conducted a subgroup to stratify menopausal status because of the rapidly increasing incidence of premenopausal breast cancer

with a unique tumor biology in East Asian patients. The present study demonstrated that estrogen-DNA adduct ratios were significantly associated with breast cancer risk in premenopausal women. The previous three studies did not identify significant interactions between estrogen-DNA adduct ratios and menopausal status in terms of risk prediction for breast cancer.

In contrast to the three earlier studies, we examined three gene polymorphisms related to estrogen metabolism. Of these, only *COMT* polymorphism exhibited a marginal association with estrogen-DNA adduct ratios. COMT catalyzes an inactivation pathway for catechol estrogen. The previous in vitro study indicated that a pretreatment of 2,3,7,8-tetrachlorodibenzo-p-dioxin followed by the inhibition of COMT activity increased the formation of depurinating 4-OHE1(E2)-1-N3Ade and 4-OHE1(E2)-1-N7Gua adducts in human breast epithelial cells [36]. With regard to the SNP, the *COMT* gene Met/Met homozygotes yielded a 3-to-4-fold reduction in COMT activity relative to the Val/Val homozygotes, and Val/Met heterozygotes demonstrated intermediate activity [37]. The frequency of the variant (low-activity) *COMT* allele was lower in Asian women than in Caucasian women (summarized in a previous article) [38]. In the present study, the estrogen-DNA adduct ratios varied in the three *COMT* SNPs (heterogeneity of the SNP groups: $P = 0.025$ using the Kruskal–Wallis test), and the Val/Met group had the highest ratio (mean ratios: Val/Val, 41.28; Val/Met, 54.32; Met/Met, 48.61). However, each of the three statistical comparisons between the two *COMT* genotypes did not reach statistical significance (data not shown). A potential explanation is that interaction of environmental pollutant exposure or lifestyle factors with decreased COMT activity, but not COMT itself, contributes to production of estrogen-DNA adducts.

To evaluate whether the formation of estrogen-DNA adducts preferentially contributes to ER-positive breast cancer, we examined the association between estrogen-DNA adduct ratios and ER/HER2 status. However, we were unable to identify a significant association. Estrogen-DNA adducts might mediate the initial steps of breast carcinogenesis in various pluripotent cell types, and therefore, estrogen-DNA adduct ratios could differentiate between the control group participants and the high breast cancer risk participants regardless of ER/HER2 status.

In summary, our study suggests that the carcinogenesis of rapidly increasing incidence of breast cancer in young women in Taiwan or other East Asian countries is related to the formation of estrogen-DNA adducts, and estrogen-DNA adduct ratios can act as biomarkers for breast cancer risk.

Abbreviations

2-OHE1(E2)
2-hydroxyestrone (estradiol)
CYP
cytochrome P450
SNP
single nucleotide polymorphisms
COMT
catechol-O-methyltransferase
National Taiwan University Hospital
SPE
solid-phase extraction
ANOVA
analysis of variance
ER
estrogen receptor
HER2 human epidermal growth factor receptor 2

Declarations

Acknowledgments

We would like to thank the participants for providing their samples and clinical information for use in our research.

Author contributions

C-H.L., W-H.K., Y-S.L., E.G.R., and A-L.C. conceived the study, and W-H.K. and M-Y.W. participated in the sample collection. M.Z., B.M., C.L.B. and E.G.R. participated in the data acquisition. F-C.H., C-H.L., and I-C.C. carried out the statistical analysis. All the authors were involved in drafting the manuscript critically for the intellectual content, and they all gave final approval for this version to be published.

Funding

This work was supported by grants from the Yonglin Foundation, the Ministry of Health and Welfare (MOHW111-TDU-B-221-114016), the Ministry of Education (NTU-109L901403), and the Ministry of Science and Technology (MOST 110-2314-B-002-214) in Taiwan.

Ethics approval and consent to participate

The procedures followed in the present study are in line with the Helsinki Declaration and have been approved by the Ethics Committee of National Taiwan University Hospital (201004060R).

Consent for publication

All the authors have read this paper and consent to its publication.

Competing interests

The authors declare that they have no competing interests.

References

1. Heinze, G., Ploner, M., Dunkler, D., Southworth, H., and Jiricka, L. (2020). logistf: Firth's bias-reduced logistic regression. R package, version 1.24 (URL: <https://CRAN.R-project.org/package=logistf>).
2. Hu, F.-C. (2017). My.stepwise: Stepwise variable selection procedures for regression analysis. R package, version 0.1.0 (URL: <https://CRAN.R-project.org/package=My.stepwise>).
3. Nagelkerke, N. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3): 691–692.
4. Yee, T. W. (2015). *Vector Generalized Linear and Additive Models: With an Implementation in R*. New York, NY: Springer Science + Business Media.
5. Yee, T. W. (2021). VGAM: Vector generalized linear and additive models. R package, version 1.1-5 (URL: <https://CRAN.R-project.org/package=VGAM>).
6. Yee, T. W. and Wild, C. J. (1996). Vector generalized additive models. *Journal of Royal Statistical Society, Series B*, 58(3): 481–493.
7. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021. doi:10.3322/caac.21660.
8. Lin CH, Yap YS, Lee KH, Im SA, Naito Y, Yeo W, et al. Contrasting Epidemiology and Clinicopathology of Female Breast Cancer in Asians vs the US Population. *J Natl Cancer Inst* 2019, 111(12):1298-1306. doi:10.1093/jnci/djz090.
9. Sim X, Ali RA, Wedren S, Goh DL, Tan CS, Reilly M, et al. Ethnic differences in the time trend of female breast cancer incidence: Singapore, 1968-2002. *BMC Cancer* 2006, 6:261.
10. Yoo KY, Kim Y, Park SK, Kang D. Lifestyle, genetic susceptibility and future trends of breast cancer in Korea. *Asian Pac J Cancer Prev* 2006, 7(4):679-682.
11. Matsuno RK, Anderson WF, Yamamoto S, Tsukuma H, Pfeiffer RM, Kobayashi K, et al. Early- and late-onset breast cancer types among women in the United States and Japan. *Cancer Epidemiol Biomarkers Prev* 2007, 16(7):1437-1442.
12. Shen YC, Chang CJ, Hsu C, Cheng CC, Chiu CF, Cheng AL. Significant difference in the trends of female breast cancer incidence between Taiwanese and Caucasian Americans: implications from age-period-cohort analysis. *Cancer Epidemiol Biomarkers Prev* 2005, 14(8):1986-1990.
13. Fan L, Strasser-Weippl K, Li JJ, St Louis J, Finkelstein DM, Yu KD, et al. Breast cancer in China. *Lancet Oncol* 2014, 15(7):e279-289. doi:10.1016/S1470-2045(13)70567-9.
14. Lin CH, Liao JY, Lu YS, Huang CS, Lee WC, Kuo KT, et al. Molecular subtypes of breast cancer emerging in young women in Taiwan: evidence for more than just westernization as a reason for the disease in Asia. *Cancer Epidemiol Biomarkers Prev* 2009, 18(6):1807-1814. doi:18/6/1807 [pii] 10.1158/1055-9965.EPI-09-0096.
15. Lin CH, Chuang PY, Chiang CJ, Lu YS, Cheng AL, Kuo WH, et al. Distinct clinicopathological features and prognosis of emerging young-female breast cancer in an East Asian country: a nationwide cancer registry-based study. *Oncologist* 2014, 19(6):583-591. doi:10.1634/theoncologist.2014-0047.
16. Lin CH, Chen YC, Chiang CJ, Lu YS, Kuo KT, Huang CS, et al. The emerging epidemic of estrogen-related cancers in young women in a developing Asian country. *Int J Cancer* 2012, 130(11):2629-2637. doi:10.1002/ijc.26249.
17. MacMahon B, Cole P, Brown JB, Aoki K, Lin TM, Morgan RW, et al. Oestrogen profiles of Asian and North American women. *Lancet* 1971, 2(7730):900-902.
18. Goldin BR, Adlercreutz H, Gorbach SL, Woods MN, Dwyer JT, Conlon T, et al. The relationship between estrogen levels and diets of Caucasian American and Oriental immigrant women. *The American journal of clinical nutrition* 1986, 44(6):945-953.
19. Shimizu H, Ross RK, Bernstein L, Pike MC, Henderson BE. Serum oestrogen levels in postmenopausal women: comparison of American whites and Japanese in Japan. *Br J Cancer* 1990, 62(3):451-453.
20. Bernstein L, Yuan JM, Ross RK, Pike MC, Hanisch R, Lobo R, et al. Serum hormone levels in pre-menopausal Chinese women in Shanghai and white women in Los Angeles: results from two breast cancer case-control studies. *Cancer Causes Control* 1990, 1(1):51-58.
21. Key TJ, Chen J, Wang DY, Pike MC, Boreham J. Sex hormones in women in rural China and in Britain. *Br J Cancer* 1990, 62(4):631-636.
22. Falk RT, Fears TR, Hoover RN, Pike MC, Wu AH, Nomura AM, et al. Does place of birth influence endogenous hormone levels in Asian-American women? *Br J Cancer* 2002, 87(1):54-60. doi:10.1038/sj.bjc.6600339.
23. Pinheiro SP, Holmes MD, Pollak MN, Barbieri RL, Hankinson SE. Racial differences in premenopausal endogenous hormones. *Cancer Epidemiol Biomarkers Prev* 2005, 14(9):2147-2153. doi:10.1158/1055-9965.EPI-04-0944.
24. Cavalieri EL, Stack DE, Devanesan PD, Todorovic R, Dwivedy I, Higginbotham S, et al. Molecular origin of cancer: catechol estrogen-3,4-quinones as endogenous tumor initiators. *Proc Natl Acad Sci U S A* 1997, 94(20):10937-10942.
25. Pruthi S, Yang L, Sandhu NP, Ingle JN, Beseler CL, Suman VJ, et al. Evaluation of serum estrogen-DNA adducts as potential biomarkers for breast cancer risk. *J Steroid Biochem Mol Biol* 2012, 132(1-2):73-79. doi:10.1016/j.jsbmb.2012.02.002.
26. Gaikwad NW, Yang L, Muti P, Meza JL, Pruthi S, Ingle JN, et al. The molecular etiology of breast cancer: evidence from biomarkers of risk. *Int J Cancer* 2008, 122(9):1949-1957. doi:10.1002/ijc.23329.

27. Dawling S, Roodi N, Mernaugh RL, Wang X, Parl FF. Catechol-O-methyltransferase (COMT)-mediated metabolism of catechol estrogens: comparison of wild-type and variant COMT isoforms. *Cancer Res* 2001, 61(18):6716-6722.
28. Dupont WD, Page DL. Risk factors for breast cancer in women with proliferative breast disease. *N Engl J Med* 1985, 312(3):146-151. doi:10.1056/NEJM198501173120303.
29. Hartmann LC, Sellers TA, Frost MH, Lingle WL, Degnim AC, Ghosh K, et al. Benign breast disease and the risk of breast cancer. *N Engl J Med* 2005, 353(3):229-237. doi:10.1056/NEJMoa044383.
30. Dyrstad SW, Yan Y, Fowler AM, Colditz GA. Breast cancer risk associated with benign breast disease: systematic review and meta-analysis. *Breast Cancer Res Treat* 2015, 149(3):569-575. doi:10.1007/s10549-014-3254-6.
31. Zahid M, Mondal B, LeVan TD, Rogan EG. Estrogen Metabolism in African-American Women with and without Breast Cancer: A Pilot Study. *Chem Res Toxicol* 2019, 32(1):190-194. doi:10.1021/acs.chemrestox.8b00285.
32. Martinez-Ramirez OC, Perez-Morales R, Castro C, Flores-Diaz A, Soto-Cruz KE, Astorga-Ramos A, et al. Polymorphisms of catechol estrogens metabolism pathway genes and breast cancer risk in Mexican women. *Breast* 2013, 22(3):335-343. doi:10.1016/j.breast.2012.08.004.
33. Yadav S, Singhal NK, Singh V, Rastogi N, Srivastava PK, Singh MP. Association of single nucleotide polymorphisms in CYP1B1 and COMT genes with breast cancer susceptibility in Indian women. *Dis Markers* 2009, 27(5):203-210. doi:10.3233/DMA-2009-0663.
34. Lajin B, Hamzeh AR, Ghabreau L, Mohamed A, Al Moustafa AE, Alachkar A. Catechol-O-methyltransferase Val 108/158 Met polymorphism and breast cancer risk: a case control study in Syria. *Breast Cancer* 2013, 20(1):62-66. doi:10.1007/s12282-011-0309-y.
35. Key T, Appleby P, Barnes I, Reeves G, Endogenous H, Breast Cancer Collaborative G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *J Natl Cancer Inst* 2002, 94(8):606-616. doi:10.1093/jnci/94.8.606.
36. Endogenous H, Breast Cancer Collaborative G, Key TJ, Appleby PN, Reeves GK, Travis RC, et al. Sex hormones and risk of breast cancer in premenopausal women: a collaborative reanalysis of individual participant data from seven prospective studies. *Lancet Oncol* 2013, 14(10):1009-1019. doi:10.1016/S1470-2045(13)70301-2.
37. Gaikwad NW, Yang L, Pruthi S, Ingle JN, Sandhu N, Rogan EG, et al. Urine biomarkers of risk in the molecular etiology of breast cancer. *Breast Cancer (Auckl)* 2009, 3:1-8. doi:10.4137/bcbcr.s2112.
38. Adlercreutz H, Gorbach SL, Goldin BR, Woods MN, Dwyer JT, Hamalainen E. Estrogen metabolism and excretion in Oriental and Caucasian women. *J Natl Cancer Inst* 1994, 86(14):1076-1082. doi:10.1093/jnci/86.14.1076.
39. Dickinson LE, MacMahon B, Cole P, Brown JB. Estrogen profiles of Oriental and Caucasian women in Hawaii. *N Engl J Med* 1974, 291(23):1211-1213. doi:10.1056/NEJM197412052912302.
40. Moore SC, Matthews CE, Ou Shu X, Yu K, Gail MH, Xu X, et al. Endogenous Estrogens, Estrogen Metabolites, and Breast Cancer Risk in Postmenopausal Chinese Women. *J Natl Cancer Inst* 2016, 108(10). doi:10.1093/jnci/djw103.
41. Visvanathan K, Yager JD. Ethnic Variations in Estrogen and Its Metabolites: Sufficient to Explain Differences in Breast Cancer Incidence Rates? *J Natl Cancer Inst* 2016, 108(11). doi:10.1093/jnci/djw223 10.1093/jnci/djw147.
42. Lu F, Zahid M, Saeed M, Cavalieri EL, Rogan EG. Estrogen metabolism and formation of estrogen-DNA adducts in estradiol-treated MCF-10F cells. The effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin induction and catechol-O-methyltransferase inhibition. *J Steroid Biochem Mol Biol* 2007, 105(1-5):150-158. doi:10.1016/j.jsbmb.2006.12.102.
43. Weinshilboum RM, Otterness DM, Szumlanski CL. Methylation pharmacogenetics: catechol O-methyltransferase, thiopurine methyltransferase, and histamine N-methyltransferase. *Annual review of pharmacology and toxicology* 1999, 39:19-52. doi:10.1146/annurev.pharmtox.39.1.19.
44. Huang CS, Lin CH, Lu YS, Shen CY. Unique features of breast cancer in Asian women—breast cancer in Taiwan as an example. *J Steroid Biochem Mol Biol* 2010, 118(4-5):300-303. doi:10.1016/j.jsbmb.2009.12.017.

Table 3

Table 3 is available in the Supplementary Files section

Figures

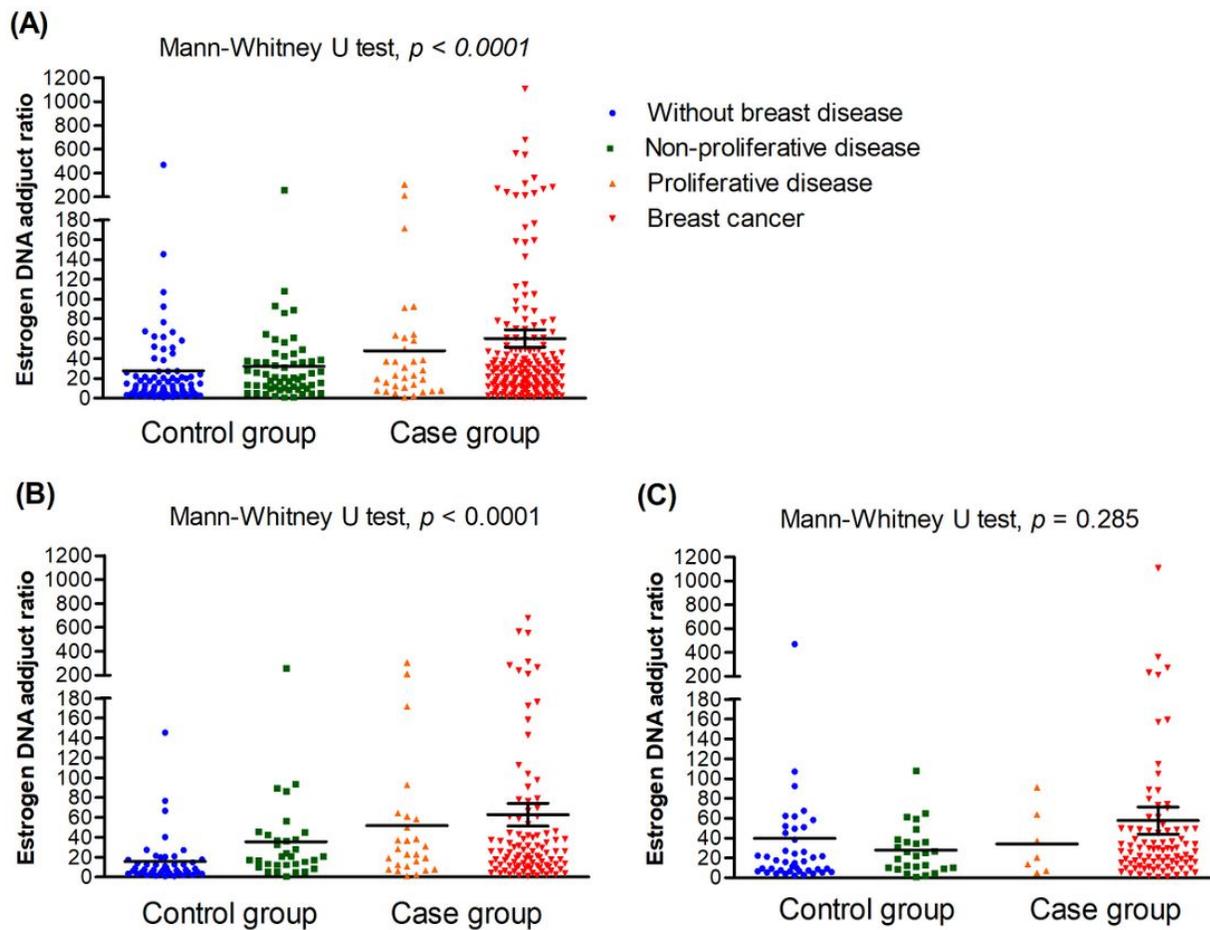


Figure 1

Ratios of estrogen-DNA adducts in women without breast disease, with non-proliferative breast disease, with proliferative breast disease, and with breast cancer for all participants (A), in the premenopausal subgroup (B), and in the postmenopausal subgroup (C).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table3.docx](#)
- [20220612SupplementaryInformation.docx](#)