Cancer Management and Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

# DNA Methylation in *RARβ* Gene as a Mediator of the Association Between Healthy Lifestyle and Breast Cancer: A Case–Control Study

This article was published in the following Dove Press journal:
*Cancer Management and Research*

Xuan Wang[1]
Yupeng Liu[1]
Hongru Sun[1]
Anqi Ge[1]
Dapeng Li[1]
Jinming Fu[1]
Yan Li[1]
Da Pang[2]
Yashuang Zhao[1]

[1]Department of Epidemiology, College of Public Health, Harbin Medical University, Harbin 150081, Heilong Jiang Province, People's Republic of China; [2]Department of Breast Surgery, The Affiliated Cancer Hospital of Harbin Medical University, Harbin 150081, Heilong Jiang Province, People's Republic of China

**Purpose:** Lifestyle factors and methylation in the retinoic acid receptor β (*RARβ*) gene are associated with breast cancer (BC). This study aims to examine the mediation effect of *RARβ* methylation on the association between healthy lifestyle and BC in Chinese women.

**Patients and Methods:** This case–control study consisted of 408 BC patients and 573 controls. A healthy lifestyle score (HLS) was constructed based on diet, alcohol use, physical activity, body mass index and smoking. The mediation effect of *RARβ* methylation in peripheral blood leukocytes was assessed in a causal mediation model using R package Lavaan.

**Results:** A higher HLS was significantly associated with lower risk of BC (*P*-value<0.001). In mediation analyses, the total effect of HLS on BC measured as a regression coefficient was significant (−0.237). The indirect effects of HLS on *RARβ* methylation (−0.153) and *RARβ* methylation on BC (0.220) were both significant. The significant mediation effect of *RARβ* methylation on the HLS-BC association was estimated at 14.3%.

**Conclusion:** The relationship between healthy lifestyle and BC is partly mediated by *RARβ* methylation, suggesting that epigenetic modifications play a role in the underlying mechanisms in response to lifestyles and contribute to the development of BC.

**Keywords:** healthy lifestyle score, retinoic acid receptor β, mediation effect, population study

## Introduction

Breast cancer (BC) is the most common cancer in women worldwide, accounting for 25.1% of all cancers.[1] Although the prevalence of BC in China is not as high as in western countries, its incidence presents a steeper upward trend than that in western societies during the last decade.[2,3] Of note, many studies have reported that adherence to healthy lifestyles significantly decreases the risk of BC;[4,5] however, the underlying mechanisms behind the association are still not well known.

Epigenetic regulation of gene expression plays an important role in the development of BC.[6] DNA methylation is one of the best-characterized epigenetic marks.[7] Studies have demonstrated that DNA methylation in some gene is associated with risk of BC.[8–10] Retinoic acid receptor β (*RARβ*) is a tumor suppressor gene, and *RARβ* methylation has been found to be associated with cancers, including BC.[11,12] Furthermore, prior studies have shown that some individual lifestyle factors, such as smoking and diet, could affect *RARβ* methylation,[13,14] although inconsistent data have been reported.[15,16] It is possible that the effect of a single lifestyle factor on the *RARβ*

Correspondence: Yashuang Zhao
Department of Epidemiology, College of Public Health, Harbin Medical University, Harbin, Heilongjiang Province 150081, People's Republic of China
Tel/Fax +86-0451-87502823
Email Zhao_yashuang@263.net

Da Pang
Department of Breast Surgery, The Affiliated Cancer Hospital of Harbin Medical University, Harbin, Heilongjiang Province 150081, People's Republic of China
Tel/Fax +86-0451-86298330
Email pangda@ems.hrbmu.edu.cn

submit your manuscript | www.dovepress.com
DovePress
http://doi.org/10.2147/CMAR.S244606

methylation is too small to be detected. By contrast, an overall lifestyle which combined multiple lifestyle factors might better evaluate the association of lifestyle factors with *RARβ* methylation.

Therefore, the existing evidence on the relationship between lifestyle behavioral factors, *RARβ* methylation and BC suggests that DNA methylation alterations in BC-related genes might be a mediator in the association between overall lifestyle and BC.

We hypothesize that *RARβ* methylation has a mediation effect on the association between healthy lifestyle and lower risk of BC. To test the hypothesis, we first evaluated the association between healthy lifestyle factors and BC risk without *RARβ* methylation included in the model. Next, we assessed the associations between healthy lifestyle score (HLS) and *RARβ* methylation and between *RARβ* methylation and BC. At last, a causal mediation analysis was conducted to assess the mediation effect of *RARβ* methylation on the healthy lifestyle–BC association in a case–control study in the Chinese woman population.[17]

# Patients and Methods

## Study Participants

We conducted a case–control study by recruiting 459 newly diagnosed BC patients (cases) in the Cancer Hospital of Harbin Medical University from 2010 to 2014. During this period, 651 cancer-free controls were recruited from the Orthopedics and Ophthalmology Departments of the Second Affiliated Hospital of Harbin Medical University. Peripheral blood (5 mL) was collected from participants on the same day of interview and stored in a −80°C freezer. Of the 459 BC patients recruited as cases, we excluded 22 patients who did not provide blood samples and 29 patients who did not complete the questionnaire. The remaining 408 BC patients were included in the case group. Of the 651 participants recruited as controls, 573 participants who provided blood samples and completed the questionnaire survey were included in the control group. Details of the recruitment process are presented in eFigure 1.

All procedures performed in this study involving human participants were carried out following the rules of the Declaration of Helsinki of 1975. The Human Research and Ethics Committee of Harbin Medical University approved this study, and all participants provided written informed consent.

A validation analysis was conducted using the European Prospective Investigation into Cancer and Nutrition (EPIC) Study cohort. A nested case–control design of the EPIC Study has been previously described.[18] The Infinium HumanMethylation450K BeadChip was used for the whole-genome DNA methylation analysis of blood buffy coat samples. DNA methylation data at four CpG sites (cg24396624, cg26786980, cg19003815 and cg27486427) in the *RARβ* gene from the epigenome-wide methylation dataset (GSE51032 from GEO) on 233 BC patients and 340 controls identified during the follow-up period were analyzed to validate the results of the *RARβ* methylation-BC association from the current study. Characteristics of the validation cohort are presented in eTable 1.

## Data Collection

The questionnaire information was obtained in a face-to-face interview. The structured questionnaire was modified based on the one previously used.[19] The Questionnaire included questions on demographic information (age, height, weight and education), menstrual history, reproductive history, family history of cancer and lifestyles (smoking, alcohol drinking, physical activity and dietary intake). Dietary intake was assessed using a validated food frequency questionnaire (FFQ) on nutrition items including beverage and foods commonly consumed in Northeast China. Frequency of alcohol drinking and food items intake in the FFQ was divided into 4–8 categories. The reference period of the questionnaire survey was one year before BC diagnosis for cases and before the interview date for controls.

## Healthy Lifestyle Score

A healthy lifestyle score (HLS) was constructed based on five lifestyle factors (diet, alcohol use, physical activity, body mass index (BMI) and smoking).[20–24] Individual lifestyle scores were defined as: 1=no current smoking, 1=no current drinking, 1=no overweight or obesity (BMI<24), 1=regular physical activity (leisure-time activities at least once per week) and 1=healthy diet. Individual diet scores were constructed by assigning 1 point for each of the following: vegetables ≥ 3 servings/day (median), fruits ≥ 3 servings/day (median), whole grains ≥ 3 servings/day (median), fish ≥ 2 servings/week (median) and red meat ≤ 1.5 servings/week (median). Healthy diet was defined as the sum of individual diet scores ≥ 3 (median). The HLS was calculated as the sum of individual scores of the five lifestyle factors (value range = 0~5). Details on the score definitions are presented in eTable 2 and eTable 3.

## DNA Extraction and Bisulfite Modification

Genomic DNA was extracted from peripheral blood leukocyte samples using a commercial DNA extraction kit (QIAamp DNA Blood Mini Kit, Hilden, Germany). DNA quantity was measured using the Nanodrop 2000 Spectrophotometer (Thermo Scientific). DNA samples were stored at −80°C until use, and then treated by bisulfite using a sodium bisulfite modification kit (EpiTect Fast DNA Bisulfite Kit, Qiagen, Hilden, Germany) according to the manufacturer's protocols. Two μg DNA was used to transform unmethylated cytosine nucleotides into thymidine without changing methylated cytosines. The yield of bisulfite-modified DNA was 50–100 ng/μL.

## *RARβ* Methylation Analysis

We performed a semi-quantitatively methylation-sensitive high-resolution melting (HRM) assay. Analysis of *RARβ* methylation was conducted on a LightCycler480 machine (Roche Applied Science, Mannheim, Germany).[25] As a candidate tumor suppressor gene, the primer pair of *RARβ* was discovered from previously published studies and further optimized using Primer Premier 5.0 software. The optimized primer sequences for HRM analysis were as follows: forward primer, 5′-CGAGTTGTTTGAGG ATTGGGATGT-3′; reverse primer, 5′-AATACGTTCC GAATCCTACCCC-3′. The amplicon (89 bp, range = chr3:25,469,838–25,469,927) was located at CpG island II in the promoter region of *RARβ* (eFigure 2).

PCR amplification system was a total of 10 ul volume consisting of 1X LightCycler480 High Resolution Melting Master Mix (Roche), 5 ng of a sodium bisulfite-modified DNA template, 200 nmol/l of each primer and 3 mmol/l of MgCl2 at final concentration. Experimental protocol of PCR amplification consisted of sufficient denaturation and activation for 10 min at 95°C for 1 cycle, denaturation for 10 s at 95°C, annealing for 30 s with a touchdown (66–56°C, 30 sec) of each primer annealing temperature and extension for 15 s at 72°C for 50 cycles. The HRM melting protocol included 95°C for 1 min, cool down to 40°C for 1 min, 68°C for 5 s and continuous acquisition to 92°C at 30 acquisitions per 1°C.

Universal methylated (100% methylated) and unmethylated (0% methylated) human whole genomic DNA samples (Zymo Research) were used as a positive and negative controls, respectively. A series of different levels of methylated standard dilutions, including 100%, 10%, 5%, 2%, 1%, 0.5% and 0%, was constructed as standard curves (eFigure 3) which were created by mixing two standards above in a corresponding ratio according to mass concentration. For quality control, DNA samples and gradient methylated DNA standards were duplicated in each plate and water was applied as blank controls.

HRM data were analyzed using Gene Scanning Software (version 2.0). Data processing included normalization and temperature shifting using a LightCycler480. The methylation status of the gene was determined by comparing the curves of each sample to the series of standard dilutions in the gene scanning module by two independent observers.

## Statistical Analysis

Analyses of covariance (generalized linear models) and Chi-square tests were used for comparison of continuous and categorical variables, respectively, between cases and controls. Adjusted odds ratios (ORs) and 95% confidence intervals (CIs) were estimated in multivariable logistic regression models to determine the associations between HLS, *RARβ* methylation and BC.

A causal mediation model (Figure 1) was constructed to examine the mediation effect of *RARβ* methylation on the HLS-BC association.[17] HLS was the predictor variable (X), *RARβ* methylation the mediator (M), and BC the outcome variable (Y). In general, there are 4 steps for mediation analyses: (1) showing that the predictor variable determines the outcome (Model Y=c X) where c is total effect; (2) showing that the predictor variable affects the mediator (Model $M = \beta_1 X$) where $\beta_1$ is indirect effect 1; (3) showing that the mediator determines the outcome controlling for the predictor (Model $Y = \beta_2 M + c' X$) where $\beta_2$ is indirect effect 2, and c' is direct effect; (4) calculating the proportion of mediation: mediation effect $(\%)=(\beta_1 \times \beta_2/c)\times 100\%$. Statistical analyses were performed using SAS 9.4 (SAS Institute Inc. Cary, NC, USA) and Lavaan package of R 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).[26]

For the analyses with BC as the dependent variable, age, education, age at menarche, parity, menopause status and family history of BC were included as covariates for adjustment.[27,28] For the indirect effect analyses of individual lifestyles and HLS on *RARβ* methylation, age and education were included as covariates for adjustment. The covariates included in the models are mentioned in footnotes of the relevant tables.

In the validation study, the associations of *RARβ* methylation levels at four CpG sites (cg24396624,

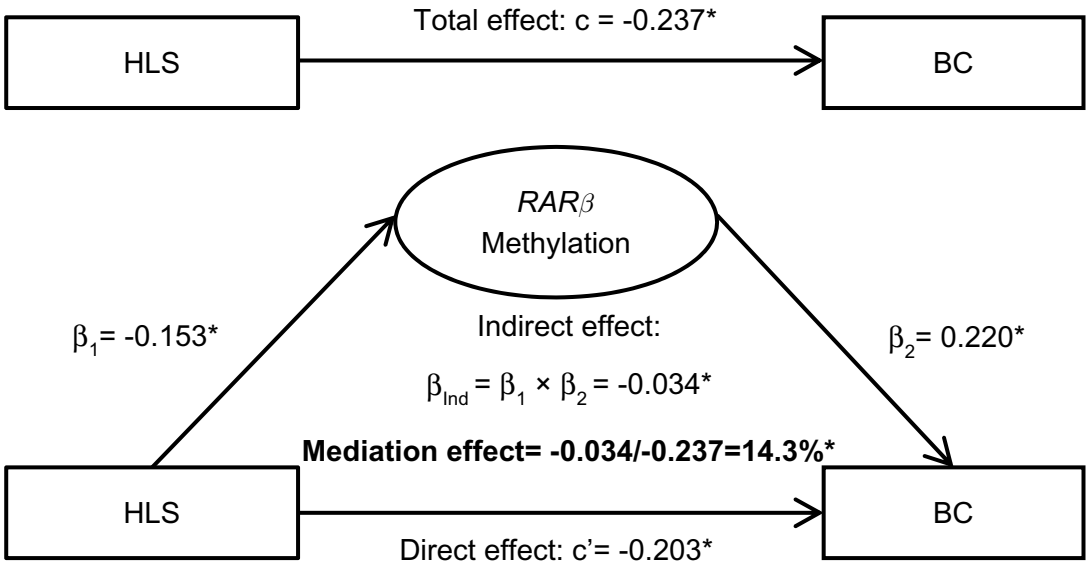


**Figure 1** Mediation Analysis Model of *RARβ* Methylation on the HLS-BC Association. β, c and c' are regression coefficients; c=total effect; c'=direct effect; $\beta_1$=indirect effect 1; $\beta_2$=indirect effect 2; $\beta_{Ind}$=total indirect effect; * indicates statistical significance at 1%.
**Abbreviations:** *RARβ*, retinal acid receptor beta; HLS, healthy lifestyle score; BC, breast cancer.

cg26786980, cg19003815 and cg27486427) with BC were examined by individual CpG sites and as a DNA methylation risk score (MRS). Two CpG sites (cg24396624 and cg26786980) that showed significant associations with BC were included in the calculation of the weighted MRS. The MRS was defined as a sum of individual methylation levels multiplied by their respective regression coefficients: weighted MRS = $\beta_1 \times CpG_1 + \beta_2 \times CpG_2$, where $\beta_i$ is the regression coefficient for $CpG_i$.

## Results

Characteristics of BC patients and controls are presented in Table 1. Age, age at menarche, and menopausal status did not show significant differences between case and control groups. Compared with controls, cases had higher levels of education, a greater number of giving birth and a higher frequency of family history of BC. Frequency of regular physical activity and healthy diet was lower in cases than in controls. HLS was significantly lower in cases than in controls. *RARβ* gene was more frequently methylated in cases than in controls.

### Association Between Healthy Lifestyle and BC

After adjustment for age, education, age at menarche, parity, menopause status and family history of BC, a higher HLS was significantly associated with lower risk of BC. The risk of BC significantly decreased with the increasing number of favorable lifestyle factors. The OR of high HLS (HLS=4~5) was 0.34 (95% CI: 0.24–0.48) for BC, with low HLS (HLS=0~2) as reference (eFigure 4).

### Association Between Healthy Lifestyle and *RARβ* Methylation

After adjusting for age and education, 1-point increase in HLS was associated with a lower risk of *RARβ* methylation (OR: 0.74, 95% CI: 0.64–0.85). For individual lifestyle factors, after additional adjustment for other lifestyle factors, healthy diet was significantly associated with a lower risk of *RARβ* methylation (OR: 0.61, 95% CI: 0.46–0.82). Other individual lifestyle factors including no current smoking, no current drinking, no overweight or obesity and regular physical activity were associated with a lower risk of *RARβ* methylation, but these associations were not significant (Table 2).

### Association Between *RARβ* Methylation and BC Risk

The associations of *RARβ* methylation with BC in the current study and validation study are shown in Table 3. After adjusting for age, education, age at menarche, parity, menopause status, and family history of BC, the *RARβ* methylation was associated with a 2.67 times higher risk of BC in the current study. *RARβ* methylation levels at cg24396624 and cg26786980 were significantly and positively associated with BC, whereas the associations were not significant at cg19003815 and cg27486427 in the

**Table 1** Characteristics of Breast Cancer Patients and Control Subjects

| Characteristics | Cases | Controls | *P*-value |
|---|---|---|---|
| **N** | 408 | 573 | |
| **Age** | 51.7 (9.3) | 51.8 (10.3) | 0.904 |
| **Age at menarche** | 15.2 (1.8) | 15.4 (2.0) | 0.118 |
| **Education, n (%)** | | | 0.030* |
| Below/at primary school | 100 (24.6) | 163 (29.0) | |
| Middle/high school | 281 (69.0) | 381 (67.8) | |
| Above/at college | 26 (6.4) | 18 (3.2) | |
| **Parity, n (%)** | | | 0.002* |
| Nullipara | 5 (1.2) | 6 (1.1) | |
| 1–2 | 203 (49.8) | 350 (61.1) | |
| ≥3 | 200 (49.0) | 217 (37.9) | |
| **Menopausal status, n (%)** | | | 0.500 |
| Premenopausal | 184 (45.1) | 246 (42.9) | |
| Postmenopausal | 224 (54.9) | 327 (57.1) | |
| **Family history of BC, n (%)** | | | 0.034* |
| No | 394 (96.6) | 565 (98.6) | |
| Yes | 14 (3.4) | 8 (1.4) | |
| **Healthy lifestyle factors, n (%)** | | | |
| No current smoking | 376 (92.2) | 542 (94.6) | 0.126 |
| No current drinking | 362 (88.7) | 528 (92.2) | 0.069 |
| No overweight or obesity | 233 (57.1) | 321 (56.0) | 0.735 |
| Regular physical activity | 138 (33.8) | 258 (45.0) | <0.001* |
| Healthy diet | 128 (31.4) | 277 (48.3) | <0.001* |
| **Health lifestyle score** | 3.01 (0.99) | 3.35 (0.93) | <0.001* |
| ***RARβ* methylation, n (%)** | | | <0.001* |
| Methylated | 172 (42.2) | 128 (22.3) | |
| Unmethylated | 236 (57.8) | 445 (77.7) | |

**Notes:** Data are presented as n (%) or mean (SD). *Indicates statistical significance at 5%
**Abbreviation:** BC, breast cancer.

validation study (eTable4). MRS was constructed using methylation values at cg24396624 and cg26786980. After adjusting for age and age at menarche, an increase of 1-point in MRS was associated with a 3.05 times higher risk of BC in the validation study (Table 3).

**Table 2** Associations of Healthy Lifestyle Score and Individual Lifestyle Factors with the Risk of *RARβ* Methylation

| Independent Variable | OR (95% CI) | *P*-value |
|---|---|---|
| HLS (Per 1-point increase) | 0.74 (0.64, 0.85)[a] | <0.001* |
| No current smoking | 0.86 (0.50, 1.49)[b] | 0.598 |
| No current drinking | 0.84 (0.53, 1.32)[b] | 0.446 |
| No overweight or obesity | 0.80 (0.61, 1.06)[b] | 0.126 |
| Regular physical activity | 0.76 (0.57, 1.03)[b] | 0.081 |
| Healthy diet | 0.61 (0.46, 0.82)[b] | 0.001* |

**Notes:** [a]Adjusted for age and education; [b]Adjusted for age, education and other lifestyle factors; *Indicates statistical significance at 5%
**Abbreviations:** *RARβ*, retinoid acid receptor beta; HLS, healthy lifestyle score; OR, odds ratio; CI, confidence interval

## Mediation Effect of *RARβ* Methylation on the Lifestyle–BC Association

Figure 1 shows the mediation effect of *RARβ* methylation on the HLS-BC association. The total effect of HLS on BC was estimated at −0.237, adjusting for age, education, age at menarche, parity, menopause status and family history of BC. The indirect effect ($\beta_{Ind}$) of HLS on BC through *RARβ* methylation was measured as the product of indirect effect 1 ($\beta_1$= −0.153) and indirect effect 2 ($\beta_2$= 0.220). The overall indirect effect ($\beta_{Ind}$) was estimated at −0.034 (−0.153×0.220). The direct effect of HLS on BC (c'= −0.203) remained significant. The percentage of the mediation effect of *RARβ* methylation was estimated at 14.3%

**Table 3** Associations of *RARβ* Methylation with the Risk of Breast Cancer in the Current Study and the Validation Study

| Independent Variable | Current Study | |
|---|---|---|
| | **OR (95% CI)** | ***P*-value** |
| *RARβ* | 2.53 (1.92, 3.35)[a]<br>2.67 (2.01, 3.55)[b] | <0.001*<br><0.001* |
| | **Validation study** | |
| | **OR (95% CI)** | ***P*-value** |
| *RARβ* MRS | 2.82 (1.99, 4.00)[a]<br>3.05 (2.12, 4.38)[c] | <0.001* |

**Notes:** MRS was constructed using *RARβ* methylation at cg24396624 and cg26786980. The greater MRS was defined as a score being higher than the median. [a]Crude OR and 95% CI; [b]Adjusted for age, education, age at menarche, parity, menopause status and family history of BC; [c]Adjusted for age and age at menarche; *Indicates statistical significance at 5%.

**Abbreviations:** *RARβ*, retinoid acid receptor beta; OR, odds ratio; CI, confidence interval; MRS, methylation risk score.

calculated as dividing the indirect effect by the total effect ($\beta_{Ind}/c = -0.034/-0.237 \times 100\%$).

## Discussion

The central findings of this case–control study are that an overall healthy lifestyle measured as the HLS was associated with lower risk of BC, and this association was partly determined by the indirect effect through *RARβ* methylation. A significant mediation effect of *RARβ* methylation on the HLS-BC association was estimated at 14.3%. These results indicate that *RARβ* methylation is one of the biological pathways linking HLS and BC at a level of epigenetic alteration.

Previous studies have demonstrated a strong association between healthy lifestyles and lower risk of BC.[4,5,29,30] However, the underlying mechanisms of the association of HLS with BC are still not very clear. A large body of evidence has shown that lifestyle factors, such as smoking and diet, could affect DNA methylation in the BC-related genes.[31,32] Furthermore, the gene methylation per se is associated with the risk of BC.[8–10] In addition to traditional possible mechanisms, such as chronic inflammation and oxidative stress, a hypothesis has been generated that lifestyles may influence BC risk through affecting epigenetic process.[33,34] To date, no studies have tested this hypothesis using the causal mediation analysis model as we applied in the current study.

Dynamics of DNA methylation is substantially influenced by metabolic processes, especially one-carbon metabolites which are closely associated with lifestyle behaviors.[35–37] *RARβ* methylation is found to be altered by lifestyle factors in some studies[13,14,38] although not all.[15,16] Genistein, a nutrient in diet, was found to be associated with *RARβ* methylation in some studies,[14,38] but not all.[16] Inconsistent results were also observed for the association between smoking and *RARβ* methylation.[13,15,39,40] A possible reason for such inconsistent results was that the effect of a single lifestyle factor on *RARβ* methylation is too small to be detected. Notably, lifestyle factors often co-exist because people follow common lifestyle patterns. Therefore, an overall indicator of multiple lifestyle factors might better evaluate the effect of lifestyles on *RARβ* methylation than individual factors. The majority of previous studies focused only on the associations between individual lifestyle factors and *RARβ* methylation. In this study, we examined the effect of HLS, an overall indicator of multiple healthy lifestyle factors, on *RARβ* methylation and BC risk. To the best of our knowledge, this is the first study to evaluate the associations of *RARβ* methylation and BC with overall lifestyle which reflect the cumulative effect of multiple lifestyle factors.

Cancer has a significant genetic and epigenetic background. *RARβ* which codes for a nuclear receptor for retinoic acid is known as a tumor suppressor gene. Methylation in the *RARβ* gene has been found to be associated with cancers, including BC.[11,12,41-43] Flanagan et al reported that alterations in *RARβ* methylation were significantly associated with BC in a case–control study.[41] Loss of *RARβ* activity may lead to abnormal cellular differentiation and suppression of apoptosis, resulting in the accumulation of aberrant cells that contributes to BC.[42,43] In this study, we detected an inverse association between *RARβ* methylation and risk of BC, and such an association was replicated in the validation study cohort.

Two systemic reviews have shown strong evidence that epigenetic mechanisms may mediate the effect of environmental factors on a wide range of human diseases.[33,34] This hypothesis was largely based on animal studies,[44,45] and the evidence from humans remains scarce. Tobi et al reported that DNA methylation in whole blood mediated the association between prenatal famine exposure and metabolic health. The mediation effects of DNA methylation at selected CpG sites were estimated from 19.6% to 28.0%.[46] Our study reinforced this epigenetic mediation hypothesis by showing a significant mediation effect of *RARβ* methylation on the association between healthy

lifestyle and BC. The mediation effect of *RARβ* methylation was 14.3% which was smaller than the direct effect of HLS on BC (85.7%). The significant direct effect of HLS on BC, independent of *RARβ* methylation, suggests that other pathophysiological, metabolic and epigenetic pathways may also mediate the lifestyle–BC association. The findings from this study indicate that combined lifestyle factors are associated with the risk of BC in part through alterations in *RARβ* methylation. The changes in *RARβ* methylation represent only one of the potential links between lifestyles and BC. Other specific causal mechanisms await elucidation.

This study had several major strengths, including the novel mediation analysis, a relatively large sample size and the use of external validation study cohort. We acknowledged that our study had certain potential limitations. Firstly, peripheral blood samples were used for *RARβ* methylation in this study. Epigenetic markers including DNA methylation profile are tissue specific.[47] Although multiple tissues such as tumor tissues are ideal for the epigenetic study of pathogenesis of BC, leukocyte DNA methylation is more feasible to investigate at a population level for obvious reasons. Studies have shown that variations in DNA methylation in peripheral blood and other tissues are correlated,[46] and DNA methylation changes in blood can be detected earlier than in target tissues.[8,9] Secondly, there might exist recall bias in the data collection process of lifestyle factors and covariates, and reverse causality regarding the causal relationship between lifestyles and BC in a case–control study design. We recruited new BC patients as cases and collected lifestyle information during the prior one year before diagnosis. This would minimize the reverse causality bias. Thirdly, we did not collect detailed information on the length of physical activity to calculate metabolic equivalents. Therefore, frequency of physical activity had to be used to define regular physical activity. However, our data showed that regular physical activity was significantly associated with lower risk of BC (OR:0.62,95% CI: 0.48–-0.81), suggesting that the current method could effectively evaluate the level of physical activity. Lastly, lifestyle data were not available in the validation cohort, and thus the association and mediation analyses involving lifestyles could not be conducted.

## Conclusion

In conclusion, our study demonstrates that an overall healthy lifestyle measured as HLS is associated with the lower risk of BC, and this association is partly mediated by *RARβ* methylation. This study provides new insights into the complex relationship between lifestyles, *RARβ* methylation and the risk of developing BC, and better understanding of the underlying epigenetic mechanisms.

## Acknowledgments

This study was funded by National Natural Science Foundation of China (grant number 81172743).

## Disclosure

Authors have no conflicts of interest.

## References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. doi:10.3322/caac.21492
2. Fan L, Strasser-Weippl K, Li JJ, et al. Breast cancer in China. *Lancet Oncol*. 2014;15:e279–289.
3. DeSantis C, Ma J, Bryan L, Jemal A. Breast cancer statistics, 2013. *CA Cancer J Clin*. 2014;64:52–62.
4. Arthur R, Wassertheil-Smoller S, Manson JE, et al. The combined association of modifiable risk factors with breast cancer risk in the women's health initiative. *Cancer Prev Res*. 2018;11:317–326.
5. McKenzie F, Ferrari P, Freisling H, et al. Healthy lifestyle and risk of breast cancer among postmenopausal women in the European prospective investigation into cancer and nutrition cohort study. *Int J Cancer*. 2015;136:2640–2648.
6. Jones PA, Baylin SB. The epigenomics of cancer. *Cell*. 2007;128:683–692.
7. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13:484–492.
8. Yang R, Pfütze K, Zucknick M, et al. DNA methylation array analyses identified breast cancer-associated HYAL2 methylation in peripheral blood. *Int J Cancer*. 2015;136:1845–1855.
9. Heyn H, Carmona FJ, Gomez A, et al. DNA methylation profiling in breast cancer discordant identical twins identifies DOK7 as novel epigenetic biomarker. *Carcinogenesis*. 2013;34:102–108.
10. Xu Z, Sandler DP, Taylor JA. Blood DNA methylation and breast cancer: a prospective case-cohort analysis in the sister study. *J Natl Cancer Inst*. 2020;112(1):87–94. doi:10.1093/jnci/djz065
11. Houle B, Rochette-Egly C, Bradley WE. Tumor-suppressive effect of the retinoic acid receptor beta in human epidermoid lung cancer cells. *Proc Natl Acad Sci U S A*. 1993;90:985–989.
12. Roman SD, Clarke CL, Hall RE, Alexander IE, Sutherland RL. Expression and regulation of retinoic acid receptors in human breast cancer cells. *Cancer Res*. 1992;52:2236–2242.
13. Vuillemenot BR, Pulling LC, Palmisano WA, Hutt JA, Belinsky SA. Carcinogen exposure differentially modulates RAR-beta promoter hypermethylation, an early and frequent event in mouse lung carcinogenesis. *Carcinogenesis*. 2004;25:623–629.
14. Fang MZ, Chen D, Sun Y, Jin Z, Christman JK, Yang CS. Reversal of hypermethylation and reactivation of p16INK4a, RARbeta, and MGMT genes by genistein and other isoflavones from soy. *Clin Cancer Res*. 2005;11:7033–7041.
15. Callahan CL, Bonner MR, Nie J, et al. Active and secondhand smoke exposure throughout life and DNA methylation in breast tumors. *Cancer Causes Control*. 2019;30:53–62.

16. King-Batoon A, Leszczynska JM, Klein CB. Modulation of gene methylation by genistein or lycopene in breast cancer cells. *Environ Mol Mutagen*. 2008;49:36–45.
17. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods*. 2010;15:309–334.
18. Cordero F, Ferrero G, Polidoro S, et al. Differentially methylated microRNAs in prediagnostic samples of subjects who developed breast cancer in the european prospective investigation into nutrition and cancer (EPIC-Italy) cohort. *Carcinogenesis*. 2015;36:1144–1153.
19. Shu XO, Yang G, Jin F, et al. Validity and reproducibility of the food frequency questionnaire used in the Shanghai women's health study. *Eur J Clin Nutr*. 2004;58:17–23.
20. Luo J, Margolis KL, Wactawski-Wende J, et al. Association of active and passive smoking with risk of breast cancer among postmenopausal women: a prospective cohort study. *BMJ*. 2011;342:d1016.
21. Duffy CM, Assaf A, Cyr M, et al. Alcohol and folate intake and breast cancer risk in the WHI observational study. *Breast Cancer Res Treat*. 2009;116:551–562.
22. Neuhouser ML, Aragaki AK, Prentice RL, et al. Overweight, obesity, and postmenopausal invasive breast cancer risk: a secondary analysis of the women's health initiative randomized clinical trials. *JAMA Oncol*. 2015;1:611–621.
23. McTiernan A, Kooperberg C, White E, et al. Recreational physical activity and the risk of breast cancer in postmenopausal women: the women's health initiative cohort study. *JAMA*. 2003;290:1331–1336.
24. Prentice RL, Caan B, Chlebowski RT, et al. Low-fat dietary pattern and risk of invasive breast cancer: the women's health initiative randomized controlled dietary modification trial. *JAMA*. 2006;295:629–642.
25. Wojdacz TK, Dobrovic A. Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. *Nucleic Acids Res*. 2007;35:e41.
26. Rosseel Y. lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48:92173.
27. Madigan MP, Ziegler RG, Benichou J, Byrne C, Hoover RN. Proportion of breast cancer cases in the United States explained by well-established risk factors. *J Natl Cancer Inst*. 1995;87:1681–1685.
28. Hiatt RA, Brody JG. Environmental determinants of breast cancer. *Annu Rev Public Health*. 2018;39:113–133.
29. Arthur R, Kirsh VA, Kreiger N, Rohan T. A healthy lifestyle index and its association with risk of breast, endometrial, and ovarian cancer among Canadian women. *Cancer Causes Control*. 2018;29:485–493.
30. McKenzie F, Ellison-Loschmann L, Jeffreys M, Firestone R, Pearce N, Romieu I. Healthy lifestyle and risk of breast cancer for indigenous and non-indigenous women in New Zealand: a case control study. *BMC Cancer*. 2014;14:12.
31. Terry MB, Delgado-Cruzata L, Vin-Raviv N, Wu HC, Santella RM. DNA methylation in white blood cells: association with risk factors in epidemiologic studies. *Epigenetics*. 2011;6:828–837.
32. Sapienza C, Issa JP. Diet, nutrition, and cancer epigenetics. *Annu Rev Nutr*. 2016;36:665–681.
33. Cortessis VK, Thomas DC, Levine AJ, et al. Environmental epigenetics: prospects for studying epigenetic mediation of exposure-response relationships. *Hum Genet*. 2012;131:1565–1589.
34. Marczylo EL, Jacobs MN, Gant TW. Environmentally induced epigenetic toxicity: potential public health concerns. *Crit Rev Toxicol*. 2016;46:676–700.
35. Campión J, Milagro FI, Martínez JA. Individuality and epigenetics in obesity. *Obes Rev*. 2009;10:383–392.
36. van Driel LM, Eijkemans MJ, de Jonge R, et al. Body mass index is an important determinant of methylation biomarkers in women of reproductive ages. *J Nutr*. 2009;139:2315–2321.
37. Monteiro JP, Wise C, Morine MJ, et al. Methylation potential associated with diet, genotype, protein, and metabolite levels in the delta obesity vitamin study. *Genes Nutr*. 2014;9:403.
38. Qin W, Zhu W, Shi H, et al. Soy isoflavones have an antiestrogenic effect and alter mammary promoter hypermethylation in healthy premenopausal women. *Nutr Cancer*. 2009;61:238–244.
39. Kiseljak-Vassiliades K, Xing M. Association of cigarette smoking with aberrant methylation of the tumor suppressor gene RARβ2 in papillary thyroid cancer. *Front Endocrinol*. 2011;2:99.
40. Ottini L, Rizzolo P, Siniscalchi E, et al. Gene promoter methylation and DNA repair capacity in monozygotic twins with discordant smoking habits. *Mutat Res Genet Toxicol Environ Mutagen*. 2015;779:57–64.
41. Flanagan JM, Munoz-Alegre M, Henderson S, et al. Gene-body hypermethylation of ATM in peripheral blood DNA of bilateral breast cancer patients. *Hum Mol Genet*. 2009;18:1332–1342.
42. Liu Y, Lee MO, Wang HG, et al. Retinoic acid receptor beta mediates the growth-inhibitory effect of retinoic acid by promoting apoptosis in human breast cancer cells. *Mol Cell Biol*. 1996;16:1138–1149.
43. Widschwendter M, Berger J, Hermann M, et al. Methylation and silencing of the retinoic acid receptor-beta2 gene in breast cancer. *J Natl Cancer Inst*. 2000;92:826–832.
44. Lin RK, Hsieh YS, Lin P, et al. The tobacco-specific carcinogen NNK induces DNA methyltransferase 1 accumulation and tumor suppressor gene hypermethylation in mice and lung cancer patients. *J Clin Invest*. 2010;120:521–532.
45. Carlin J, George R, Reyes TM. Methyl donor supplementation blocks the adverse effects of maternal high fat diet on offspring physiology. *PLoS One*. 2013;8:e63549.
46. Tobi EW, Slieker RC, Luijk R, et al. DNA methylation as a mediator of the association between prenatal adversity and risk factors for metabolic disease in adulthood. *Sci Adv*. 2018;4:eaao4364.
47. Slieker RC, Bos SD, Goeman JJ, et al. Identification and systematic annotation of tissue-specific differentially methylated regions using the Illumina 450k array. *Epigenetics Chromatin*. 2013;6:26.