HAL
open science

# Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis

Zhong-Zheng Zhu, L Hou, Valentina Bollati, Letizia Tarantini, Barbara Marinelli, Laura Cantone, Andrea Baccarelli

## To cite this version:

Zhong-Zheng Zhu, L Hou, Valentina Bollati, Letizia Tarantini, Barbara Marinelli, et al.. Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis. International Journal of Epidemiology, 2010, 10.1093/ije/DYQ154 . hal-00623791

**HAL Id: hal-00623791**

**https://hal.science/hal-00623791v1**

Submitted on 15 Sep 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OXFORD UNIVERSITY PRESS
International Journal of Epidemiology

# Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis

| | |
|---|---|
| Journal: | *International Journal of Epidemiology* |
| Manuscript ID: | IJE-2009-11-1067.R2 |
| Manuscript Type: | Original Article |
| Date Submitted by the Author: | 24-Jul-2010 |
| Complete List of Authors: | Zhu, Zhong-Zheng; Università degli Studi di Milano and IRCCS Maggiore Policlinico Hospital, Mangiagalli and Regina Elena Foundation, Center of Molecular and Genetic Epidemiology, Department of Environmental and Occupational Health; No.3 People's Hospital, School of Medicine, Shanghai Jiaotong University, Department of Oncology<br>Hou, L; Feinberg School of Medicine, Northwestern University, Department of Preventive Medicine & Robert H. Lurie Comprehensive Cancer Center<br>Bollati, Valentina; Università degli Studi di Milano and IRCCS Maggiore Policlinico Hospital, Mangiagalli and Regina Elena Foundation, Center of Molecular and Genetic Epidemiology, Department of Environmental and Occupational Health<br>Tarantini, Letizia; Università degli Studi di Milano and IRCCS Maggiore Policlinico Hospital, Mangiagalli and Regina Elena Foundation, Center of Molecular and Genetic Epidemiology, Department of Environmental and Occupational Health<br>Marinelli, Barbara; Università degli Studi di Milano and IRCCS Maggiore Policlinico Hospital, Mangiagalli and Regina Elena Foundation, Center of Molecular and Genetic Epidemiology, Department of Environmental and Occupational Health<br>Cantone, Laura; Università degli Studi di Milano and IRCCS Maggiore Policlinico Hospital, Mangiagalli and Regina Elena Foundation, Center of Molecular and Genetic Epidemiology, Department of Environmental and Occupational Health<br>Baccarelli, Andrea; University of Milan & IRCCS Maggiore Hospital Mangiagalli and Regina Elena Foundation, Centre of Molecular and Genetic Epidemiology |
| Key Words: | Blood, DNA methylation, epigenetics, meta-analysis, repetitive elements |
| | |

SCHOLARONE™
Manuscripts

For Review Only

# Predictors of global methylation levels in blood DNA of healthy subjects: a combined analysis

Zhong-Zheng Zhu,[1,2§] Lifang Hou,[3§] Valentina Bollati,[1] Letizia Tarantini,[1] Barbara Marinelli,[1] Laura Cantone,[1] Allen S. Yang,[4] Pantel Vokonas,[5] Jolanta Lissowska,[6] Silvia Fustinoni,[1] Angela C. Pesatori,[1] Matteo Bonzini,[7] Pietro Apostoli,[8] Giovanni Costa,[1] Pier Alberto Bertazzi,[1] Wong-Ho Chow,[9] Joel Schwartz[10] and Andrea Baccarelli[1,10]*

[1] Center of Molecular and Genetic Epidemiology, Department of Environmental and Occupational Health, Università degli Studi di Milano and IRCCS Maggiore Policlinico Hospital, Mangiagalli and Regina Elena Foundation, Milan, Italy

[2] Department of Oncology, No.3 People's Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China

[3] Department of Preventive Medicine & Robert H. Lurie Comprehensive Cancer Center, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

[4] Division of Hematology, Keck School of Medicine, University of Southern California, Norris Comprehensive Cancer Center, Los Angeles, California, USA

[5] VA Boston Healthcare System, Department of Medicine, Boston University School of Medicine, Boston, MA, USA

[6] Department of Cancer Epidemiology and Prevention, The M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland

[7] Department of Clinical and Biological Sciences, University of Insubria, Varese, Italy

[8] Department of Experimental and Applied Medicine, Occupational Medicine and Industrial Hygiene, University of Brescia, Brescia, Italy

[9] Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, Rockville, MD, USA

[10] Exposure Epidemiology and Risk Program, Department of Environmental Health, Harvard School of Public Health, Boston, MA, USA

[§]Drs. Zhu and Hou contributed equally to this work

*Corresponding Author. Center of Molecular and Genetic Epidemiology, Department of Environmental and Occupational Health, Università degli Studi di Milano, Via San Barnaba 8, 20122 Milan, Italy, Tel: +39 02 503 20145, Fax: +39 02 503 20103, Email: andrea.baccarelli@unimi.it

Word count for paper plus abstract: 6023.

For Review Only

## Summary

**Background** Estimates of global DNA methylation from repetitive DNA elements, such as Alu and LINE-1, have been increasingly used in epidemiology investigations because of their relative low cost, high throughput, and quantitative results. Nevertheless, determinants of these methylation measures in healthy individuals are still largely unknown. The aim of this study was to examine whether age, gender, smoking habits, alcohol drinking, and body mass index (BMI) are associated with Alu or LINE-1 methylation levels in blood leukocyte DNA of healthy individuals.

**Methods** Individual data from five studies including a total 1465 healthy subjects were combined. DNA methylation was quantified by PCR-pyrosequencing.

**Results** Age ($\beta$ = −0.011 %5mC/year, 95% CI −0.020 to −0.001 %5mC/year) and alcohol drinking ($\beta$ = −0.214, 95% CI −0.415 to −0.013) were inversely associated with Alu methylation. Compared with females, males had lower Alu methylation ($\beta$ = −0.385, 95% CI −0.665 to −0.104), and higher LINE-1 methylation ($\beta$ = 0.796, 95% CI 0.261 to 1.330). No associations were found with smoking or BMI. Percent neutrophils and lymphocytes in blood counts exhibited a positive ($\beta$ = 0.036, 95% CI 0.010 to 0.061) and negative ($\beta$ = −0.038, 95% CI −0.065 to −0.012) association with LINE-1 methylation, respectively.

**Conclusions** Global methylation measures in blood DNA vary in relation with certain host and lifestyle characteristics, including age, gender, alcohol drinking, and white-blood-cell counts. These findings need to be considered in designing epidemiological investigations aimed at identifying associations between DNA methylation and health outcomes.


**Keywords** Blood, DNA methylation, epigenetics, meta-analysis, repetitive elements

## Introduction

DNA methylation is a well-defined epigenetic mechanism involved in the regulation of a wide variety of biological processes, including gene expression, genomic stability, and parental imprinting.[1] In mammals, methylation involves addition of methyl groups to cytosine to form 5-methyl-cytosine (5mC). It is estimated that more than one-third of DNA methylation occurs in repetitive elements,[2,3] which represent a large portion of the human genome.[4] Among these repetitive sequences, Alu and LINE-1 are the most plentiful families representing approximately 30% of the human genome.[5-7] Because of their high representation throughout the genome, Alu and LINE-1 have been used as surrogate markers for estimating global DNA methylation level.[8,9] Methodologically, a high-throughput quantitative method based on bisulfite treatment of DNA and subsequent PCR pyrosequencing[9] or real-time PCR[8] has been developed for estimating global DNA methylation level of Alu and LINE-1 elements. This method is amenable to use in screening large sets of samples because it is cost-effective and requires relatively small amounts of genomic DNA. As a result, DNA methylation analysis of Alu and/or LINE-1 methylation has been used in a number of epidemiological investigations.[10-27] Because of the growing interest in DNA methylation, as well as the wide availability of the technology necessary to conduct DNA methylation analyses, the number of epidemiological investigations including DNA methylation measures will increase rapidly in the next few years. Information about how DNA methylation measures vary in relation with subjects' characteristics and other information commonly collected in epidemiological investigations will be relevant for effective design and statistical analysis of such studies.

Abnormal DNA methylation patterns have been associated with various human diseases, including cancer, cardiovascular disease and autoimmune disease.[28-30] In healthy populations, inter-individual methylation variations at Alu and LINE-1 elements from blood DNA have

For Review Only

also been associated with risk factors for cancer, neurological and cardiovascular diseases.[10-13] As recently discussed by Foley *et al.*,[31] the identification and validation of epigenetic markers measured in easily obtainable tissues will be required for epidemiologic studies to benefit from incorporation of epigenetics. Among the biospecimens suitable for DNA methylation analysis, blood DNA has been the most widely collected in epidemiology investigations. Although global DNA methylation levels have been proposed to be influenced by a number of endogenous and exogenous parameters,[32] determinants of blood DNA methylation in repetitive elements in healthy individuals are still largely unexplored.

Although a number of demographic and lifestyle factors have been associated with global hypomethylation, the results have been inconsistent including those from our own studies.[12-14] Because of the moderate inter-individual differences in global methylation measures in blood leukocyte DNA, the limited sample size of each investigation may have contributed to the inconsistencies. In a study of elderly individuals in the VA Normative Aging Study (NAS), we have shown that repetitive element methylation was cross-sectionally associated with age, particularly in Alu element.[14] In the present work on a combined analysis of 1465 subjects from five investigations, we sought to determine the association of age, gender, smoking habits, alcohol drinking, and body mass index (BMI) with the methylation status of Alu and LINE-1 elements among healthy individuals. Because DNA methylation was analyzed on DNA from white blood cells in all studies, whether differences in white blood cell count were associated with Alu and LINE-1 methylation was also explored.

## Methods

### Study populations

Individual data from five studies were combined including a total 1465 healthy subjects: [1] 722 subjects from the VA Normative Aging Study (NAS) in Eastern Massachusetts, USA; [2]

For Review Only

421 subjects from controls of a population-based case-control study of gastric cancer conducted in Warsaw, Poland; [3] 211 subjects from a cross-sectional study of low-benzene exposure conducted in Milan, Italy; [4] 63 healthy workers from a steel plant in Brescia, Italy; and [5] 48 healthy workers from a perfluooctanoic acid (PFOA) production facility in Trissino, Italy. All participants in studies #2, 3, 4 and 5, and 98.2% in study #1 were Caucasians. Detailed information about gender, age, height, weight, drinking and smoking was obtained by structured questionnaires, either self-administered (studies #1, 3 and 4) or through in-person interviews (studies #2 and 5). White blood cell counts were performed on fresh blood using automated methods (Coulter Cell Counter, Abbott or Bayer) within 3 hours from the time of blood drawing for all studies. Informed consent was obtained from all participants and the ethical committees of all participating centers approved the studies. Details on each of the studies are provided below.

Study #1 included 722 men who were evaluated between March 1999 and October 2007 as part of the Normative Aging Study (NAS), a longitudinal study of normal aging established in 1963 by the U.S. Veterans Administration.[33] The men, aged 55 to 100 during 1999-2007, are given a health examination every 3-5 years and asked to donate a 7 ml blood sample at each visit. DNA methylation data from only the first blood sample of each study participant was included in our analysis.

Study #2 was a population-based case-control study of gastric cancer conducted in Warsaw, Poland during 1994-1996.[34] Four hundred and twenty-one subjects included in the present analysis were non-cancer controls randomly selected from a computerized registry of the Warsaw population, frequency-matched by age and gender to the cases.

Study #3 was carried out among residents of Milan, Italy, from October 1999 to June 2000.[12] This study included 77 gasoline station attendants, 77 traffic-control officers (both

For Review Only

groups exposed to low-benzene levels), and 57 office workers. All subjects were free of major illnesses.

Study #4 included 63 healthy male workers from a steel production plant in Brescia, Italy.[17] Between April and May 2006, every subject was asked to donate a blood sample at two different times. In the current study, only blood samples of every participant collected in the morning of the first work week day (after 2 days off work) before the beginning of any work activity were analyzed.

Study #5 included 48 male workers who were employed in PFOA production during 1978-2007.[35] Blood samples were collected in 2007, when 35 were still active workers in the PFOA department, whereas 13 were no longer exposed after retirement or being transferred to other departments.

**Completeness of the data**

Some of the variables evaluated in our analysis were not available in all studies. BMI of the 211 subjects from Study #3 and information on alcohol drinking about the 63 subjects from Study #4 were not recorded. Detailed information on smoking except for smoking status (nonsmoker, former and current smoker) was not collected from the 48 subjects from Study #5. In addition, white blood cell count data were not available on the 421 subjects from Study #2, and only data on lymphocyte and monocyte counts, but not on neutrophils, eosinophils, and basophils, were available on the 63 subjects from Study #4.

**Definition of the variables of interest**

All variables used in the present analysis were defined for all studies using the same criteria. An ever-smoker was defined as a smoker of at least 1 cigarette per day for 6 months or longer. Pack-years of smoking were calculated as the product of packs of cigarettes smoked per day and total years of smoking. An ever-drinker was defined as a person who reported drinking

For Review Only

alcoholic beverage at least 1 time per month for 6 months or longer. The age and BMI at blood draw were used for all analyses.

**Alu and LINE-1 methylation analyses**

Alu and LINE-1 methylation analyses were conducted using the method developed by Yang *et al*.[9] for all the studies, as previously described.[12,14,17] One of the study collaborators (V.B.) participated in all DNA methylation analyses, either by directly performing them or supervising the laboratory protocol application. Blood leukocyte DNA was extracted, bisulfite treated, and PCR amplified. Pyrosequencing was done using the PSQHS96 Pyrosequencing System (Pyrosequencing). The degree of methylation was expressed for both Alu and LINE-1 as percentage of methylated cytosines (%5mC). For all assays built-in controls were used to verify bisulfite conversion. Each marker was tested in at least two replicates and their average was used in the statistical analysis. The within-sample coefficient variations in duplicate runs were 7.0% and 5.3% for Alu and LINE-1, respectively. Samples were dropped from analysis when replication failed or replicates were not within a specified range (5%). Pyrosequencing data for each marker were successful obtained for ≥98.6% of the samples.

**Statistical methods**

Multivariable linear regression models were used to evaluate Alu and LINE-1 methylation levels in relation to i) subjects' characteristics (age, gender, BMI, smoking status, current cigarettes/day, past cigarettes/day, total smoking years, pack-years, age smoking started, age smoking stopped and alcohol drinking) with adjustment for age (as a continuous variable), gender (male vs. female), and study (using indicator variables), and to ii) total number of white blood cells and the proportions of white-blood-cell types (neutrophils, lymphocytes, monocytes, eosinophils and basophils) with adjustment for age (as a continuous variable), gender (male vs. female), alcohol drinking (ever-drinker vs. nondrinker) and study (using

For Review Only

indicator variables). To evaluate whether differences in differential blood counts confounded the potential associations between subjects' characteristics and DNA methylation, multivariable models were fitted adjusting for percent lymphocytes, in addition to age, gender and study. Since blood count data were missing in Study #2 (Warsaw, Poland), percent lymphocyte data were imputed using linear prediction from a multiple-equation (ME) regression model[36] that included age, gender, smoking status, current cigarettes/day, pack-years and alcohol drinking. All statistical analyses were conducted using Stata 10.1 (Stata Corporation, College Station, TX).

For Review

**Results**

The characteristics of the study populations, as well as of the combined data, are shown in Table 1. In the combined data, mean blood DNA methylation levels, expressed as percentage of 5-methylcytosine (%5mC), were 26.0 (SD = 1.8) for Alu, and 76.2 (SD = 6.2) for LINE-1.

The associations between the subjects' characteristics and Alu and LINE-1 methylation are shown in combination, as well as within each of the studies (Table 2). In the combined data, age ($\beta$ = −0.011 %5mC/year, 95% CI −0.020 to −0.001 %5mC/year) was negatively associated with DNA methylation in Alu elements. Compared with females, males had a lower Alu methylation levels ($\beta$ = −0.385, 95% CI −0.665 to −0.104), but a higher LINE-1 methylation levels ($\beta$ = 0.796, 95% CI 0.261 to 1.330). As Studies #1, 4 and 5 contained males only, further analyses of gender in relation to Alu and LINE-1 methylation were repeated using only subjects from Studies #2 and 3. In this analysis (Supplementary Table S1), no major differences in effect estimates were observed from the results reported above. In addition, alcohol drinkers tended to have a lower Alu methylation levels than nondrinkers ($\beta$ = −0.214, 95% CI −0.415 to −0.013). Age and alcohol drinking showed no association with LINE-1 methylation level. DNA methylation in both Alu and LINE-1 elements was not

associated with BMI, smoking status (Table 2), current cigarettes/day, past cigarettes/day, total smoking years, pack-years, age smoking started or age smoking stopped (data not shown). Further analyses on the associations of age, gender, and BMI with Alu and LINE-1 methylation were conducted on subsets of the study population by including alcohol drinkers only (n = 246), smokers only (n = 371), or individuals who were both alcohol drinkers and smokers (n = 481). These sets of analyses produced comparable results (data not shown). As smoking habits, alcohol drinking, and BMI differed between males and females ($P < 0.001$ for all, Table 1), we also performed gender-stratified analyses of smoking habits, alcohol drinking, and BMI, which showed no major differences in their associations with Alu or LINE-1 methylation (data not shown). Since 265 subjects (18.1% of the total number of subjects) were exposed to certain occupational pollutants, including benzene,[12] particulate matter[17] and PFOA,[35] which may exert some effects on global DNA methylation and thus modify the effects of characteristics on global DNA methylation, further association analyses of subjects' characteristics with DNA methylation were repeated by excluding these exposed subjects. In this analysis (Supplementary Table S2), only minor departures of the effect estimates were observed from the results based on the total number of subjects.

As DNA methylation was measured on white blood cells, whether the differential white blood cell count was associated with Alu and LINE-1 element methylation was examined (Table 3). The total number of white blood cells was not associated with either Alu or LINE-1 methylation. Percent neutrophils and lymphocytes, the two major types of white blood cells, were inversely correlated with each other ($r = -0.936$, $P < 0.001$) and exhibited a positive ($\beta = 0.036$, 95% CI 0.010 to 0.061) and negative ($\beta = -0.038$, 95% CI −0.065 to −0.012) association with LINE-1 methylation, respectively. No association between the differential blood count and Alu element methylation was found. Adding percent lymphocytes into multivariable linear regression models produced no major difference in the effects of

characteristics on Alu and LINE-1 methylation level (Table 4). When these analyses were repeated without imputation of percent lymphocytes for Study #2 (i.e., by excluding Study #2), the results were generally similar to the ones reported (data not shown). The only noticeable difference was that alcohol drinking was not associated with Alu methylation in this set of models ($\beta = -0.178$, 95% CI −0.433 to 0.076). However, this appeared to result from the exclusion of Study #2, which was the study with the highest number of ever drinkers (n=283), rather than from adjustment for percent lymphocytes, as also demonstrated in a model adjusting for age, gender and study, but not for percent lymphocytes, fitted without including Study #2 ($\beta$ for alcohol drinking = −0.177, 95% CI −0.428 to 0.073).

## Discussion

In a combined analysis of blood DNA methylation from five investigations, we showed that age, gender and alcohol drinking are predictors for global methylation levels in blood DNA from healthy subjects. In particular, age and alcohol were inversely associated with Alu element methylation. Gender had opposite associations with Alu and LINE-1 methylation levels, with males having lower Alu and higher LINE-1 methylation levels than females. We also showed that the neutrophil and lymphocyte proportions of white-blood-cell types modified LINE-1 methylation level measured in blood DNA, but did not confound the association between global methylation levels and subjects' age, gender, and alcohol drinking. Although Alu and LINE-1 methylation levels were shown to be correlated with each other in tumor tissue DNA,[26] no such correlation has been found in surrounding normal tissues,[26] or in non-target tissues such as blood and buccal cells.[12,13,37] In this respect, the differences between Alu and LINE-1 in methylation regulation mechanisms,[38] responses to cellular stressors and environmental exposures,[12,13,37] and average levels of DNA methylation[9,13] may account for the different findings for Alu and LINE-1 methylation in association with the

For Review Only

potential determinants evaluated in the present study, such as age, gender, alcohol drinking and the proportions of white-blood-cell types.

The effect of age on global DNA methylation has been widely discussed.[39,40] In analysis of blood DNA derived from healthy subjects, several studies based on relatively small study samples (between 76 and 237 subjects) reported an inverse association between age and genomic 5mC content,[41,42] concluding that age is an independent predictor for genomic 5mC content. However, several other similarly sized or larger studies (between 32 and 526 samples) reported no association of age with genomic 5mC content[43] or other estimates of global DNA methylation obtained from various genomic components, including LINE-1, Alu, and CCGG sequences.[13,44-46] Our finding that Alu, but not LINE-1, methylation levels decreased in association with age suggests that age-related decreases in DNA methylation are not occurring homogenously across the human genome. Also, our recent longitudinal analysis in the VA Normative Aging Study,[14] along with analyses in the present study and in a previous investigation,[41] showed that age-related methylation alteration are small, particularly compared with the inter-individual variability in DNA methylation. Thus, the lack of association with age in several studies[13,44-46] could be partially due to smaller sample size, insufficient age-range evaluated, and/or limitations of the analytical methods used.

We found that gender had opposite associations with Alu and LINE-1 methylation levels, with males having lower Alu but higher LINE-1 methylation levels. Previous studies have also observed that males, compared with females, had elevated LINE-1 methylation in blood DNA,[15,45] consistent with our data for LINE-1. Yet, our observation that males had lower Alu methylation levels may seem paradoxical. One of the two X chromosomes in females is transcriptionally inactive and heavily methylated. This, together with differences in Alu element density and composition between chromosome X and Y,[5] may partly account for the lower Alu methylation levels observed in males. Two investigations of global 5mC content,

For Review Only

reflecting a net effect of increase and decrease of methylation of various components in genomic DNA,[47] revealed a higher global 5mC content of blood DNA in males than in females.[41,42] Our data seemed to be consistent with these findings, since male gender was associated with higher $\beta$ estimates for LINE-1 ($\beta = 0.796$) than for Alu ($\beta = -0.385$), which may contribute to an elevated net global methylation levels in males.

Smoking has been associated with a state of global hypomethylation in genomic DNA derived from tumor tissues in head and neck squamous cell carcinoma (HNSCC).[19,45,48,49] Vaissiere *et al.*[50] found that smoking was associated with increased methylation in lung cancer tissues of MTHFR, a gene participating in the methyl donor pathway, and that in turn MTHFR methylation showed a marked inverse correlation with LINE-1 methylation. We hypothesized that smoking would exert effects on global methylation levels in genomic DNA derived from blood of healthy subjects, but failed to find any association between smoking habits (smoking status, current cigarettes/day, past cigarettes/day, total smoking years, pack-years, age smoking started and age smoking stopped) and Alu or LINE-1 methylation levels. Consistent with our data, other studies of smoking effects on global methylation in blood DNA of healthy subjects also reported negative results.[13,43,45] The lack of association between smoking habits and global methylation in genomic DNA from blood of healthy subjects, which are at variance with inverse associations found in tumor tissues,[19,45,48,49] may suggest that smoking acts differently in different tissue types, or acts differently in healthy versus tumor tissues. Although we did not observe a direct effect of smoking on global DNA methylation among healthy subjects, a possible involvement of smoking in modifying global DNA methylation cannot be excluded. Hillemacher *et al.*[51] recently reported that smoking modifies the correlation between paternal and offspring DNA methylation. Although neither paternal smoking nor offspring smoking had effect on global DNA methylation estimated in CCGG sequences, correlation between offspring and paternal DNA methylation was strongest

For Review Only

when both were non-smokers, and absent when the offsprings smoked.[51] In addition, Breton *et al*.[52] reported that prenatal exposure to tobacco smoke was associated with decreased global methylation level in buccal cell DNA in children. Because DNA methylation is largely established in utero or in early life,[53] these findings may reflect higher susceptibility to smoking in specific periods of life.

Several studies have examined the association between alcohol drinking and global 5mC content or LINE-1 element methylation in both normal and tumor tissue DNA, with most studies reporting no association.[43,45,48,49] Only one study of HNSCC tumor tissues reported that alcohol use was inversely associated with LINE-1 methylation levels in univariate analysis, but this significance disappeared in multivariate analysis.[19] In our data, no association was found between alcohol drinking and LINE-1 methylation, while we showed for the first time that drinking was inversely associated with Alu methylation levels. Alcohol drinkers might have lower intakes as well as lower serum levels of vitamin $B_{12}$ and folate,[54] and thus have lower availability of methyl sources, as well as decreased efficiency of the enzymatic pathways responsible for DNA methylation.[54,55] In addition, animal models have shown that chronic alcohol exposure can decrease cytosine methyltransferase mRNA levels in sperm[56] and DNA methylation in tissues of liver and colon mucosa.[57] It should be noted that most of the studies included in the present analysis did not have detailed data on alcohol drinking, which hindered our ability to further evaluate alcohol dose-related effects on global DNA methylation. Moreover, our present analysis, together with all other studies on the association between alcohol drinking and global DNA methylation,[19,43,45,48,49] mostly included social drinkers. Additional studies on alcohol abusers and dependence are required.

Our data showed that the proportions of white-blood-cell subsets affected LINE-1 methylation levels measured in blood DNA, with percent neutrophils and lymphocytes, the major types of white blood cells, exhibiting opposite effects on LINE-1 methylation levels.

For Review Only

Previous studies have shown that there exist between-tissue methylation variations, which contribute to regulating functions that are specific to each tissue or cell types.[58] In the present work, further analysis showed that the addition in multivariable regression models of percent lymphocytes did not modify the associations of the subjects' characteristics we evaluated with repetitive-element methylation levels. However, the association we observed between differential white blood count and LINE-1 methylation indicates that findings from epidemiological studies based on DNA methylation analyses on blood DNA might be confounded by differences in the proportion of white blood cell types. This would be a concern for exposures or disease states that are associated with inflammatory responses that may produce shifts in white blood cell type proportions.[59,60]

Our study had the advantages of being based on a large sample size and accurate quantitative analysis using pyrosequencing methodology, which is suitable for measuring subtle changes in DNA methylation. In the present work, all DNA methylation analyses were performed using the same laboratory protocol, and each variable used in the present analysis was defined using the same criteria. To account for potential confounding due to differences between the five studies, all our analyses were adjusted by study using indicator variables in the regression models, thus also minimizing the potential effect on methylation caused by between-study differences in ethnic background, which has been shown to be associated with global DNA methylation levels,[18,61] as well as potential analytical batch effects. One limitation of our analysis is that results on the association of gender with LINE-1 or Alu methylation could be confounded by the origin of the study and by the difference in lifestyle factors between males and females. Only two (Studies #2 and 3) of the five studies used in the combined analysis included females (only 14.1% of the study subjects were females). In addition, smoking habits, alcohol drinking, and BMI differed between males and females. However, similar effect estimates were observed after excluding studies containing males

For Review Only

only (Studies #1, 4 and 5) from the analyses (Supplementary Table S1), or by further adjusting for variables in multivariable analyses that might have confounded the associations, including smoking habits, alcohol drinking, BMI and percent lymphocytes (data not shown). Nonetheless, it cannot be excluded that other unmeasured lifestyle or personal factors might underline the gender-related differences in DNA methylation levels we observed. Secondly, dietary differences were not investigated here, which may also influence methylation levels in genomic DNA.[62] Lastly, we assumed that DNA methylation in Alu and LINE-1 repetitive elements may be used as a surrogate for global DNA methylation content. However, methylation levels of both repetitive elements are not equivalent to, albeit vastly represents, global DNA methylation content.[9] Also, measures of repetitive element-based methylation analyses as used in the present study, representing a pool of multiple unique Alu or LINE-1 repeats in the entire genome, are different from measures of gene-specific or genome-wide array-based (genomic site-specific) methylation analyses, and thus the extent to which our conclusions apply to site-specific methylation measures remains to be determined. In particular, whereas the present of associations of the variables we evaluated with Alu and/or LINE-1 methylation might suggest that there may also exist associations with gene-specific methylation, lack of association with Alu and LINE-1 methylation does not exclude associations with specific loci. Both Alu and LINE-1 methylation levels represent an average measure of methylation at multiple sites across the genome. Opposite concurrent changes of hypo- and hyper-methylation might offset each other without being reflected in these average measures of methylation.

In conclusion, this combined analysis of blood DNA methylation showed that study participant data that are routinely collected in most epidemiological investigations, including age, gender and alcohol drinking, are predictors of global DNA methylation levels. We also showed that the neutrophil and lymphocyte proportions of white-blood-cell types modified

For Review Only

LINE-1 methylation level measured in blood DNA. These findings provide valuable information that can be used in planning future epidemiological studies, as well as in analyzing existing datasets.

**Funding**

This work was supported by grants as follows:

Study #1: National Institute of Environmental Health Sciences [ES015172-01]; and Environmental Protection Agency [R83241601, R827353]. The VA Normative Aging Study is supported by the Cooperative Studies Program/Epidemiology Research and Information Center of the U.S. Department of Veterans Affairs and is a component of the Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC).

Study #2: Intramural Research Program (Division of Cancer Epidemiology and Genetics, NCI) of the National Institutes of Health; Diane Belfer Program for Human Microbial Diversity; and the National Institutes of Health [R01GM63270].

Study #3: Italian Association for Cancer Research (AIRC) (2008-2010) [6016].

Study #4: Cariplo Foundation [2007-5469]; and Italian Ministry of Scientific Research (MIUR) [PRIN-20072S2HT8].

Study #5: Milan University intramural funding.

**Conflict of interest:** None declared.

For Review Only

**KEY MESSAGES**

- Information on how global methylation measured in blood leukocyte DNA varies in relation with subjects' characteristics and other factors commonly collected in epidemiological investigations will be relevant for effective study design and statistical analysis.
- We showed that study participant data that are generally collected in most epidemiological investigations, including age, gender, alcohol drinking, and blood neutrophil and lymphocyte proportions, are predictors of global DNA methylation levels.
- These findings need to be considered in designing epidemiological investigations aiming at identifying associations between DNA methylation and health outcomes.

iew Only

## References


[1] Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 2002;**16**:6-21.

[2] Kochanek S, Renz D, Doerfler W. DNA methylation in the Alu sequences of diploid and haploid primary human cells. *Embo J* 1993;**12**:1141-51.

[3] Schmid CW. Does SINE evolution preclude Alu function? *Nucleic Acids Res* 1998;**26**:4541-50.

[4] Lander ES, Linton LM, Birren B *et al.* Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860-921.

[5] Grover D, Mukerji M, Bhatnagar P, Kannan K, Brahmachari SK. Alu repeat analysis in the complete human genome: trends and variations with respect to genomic composition. *Bioinformatics* 2004;**20**:813-7.

[6] Babushok DV, Kazazian HH Jr. Progress in understanding the biology of the human mutagen LINE-1. *Hum Mutat* 2007;**28**:527-39.

[7] Kazazian HH Jr, Goodier JL. LINE drive: retrotransposition and genome instability. *Cell* 2002;**110**:277-80.

[8] Weisenberger DJ, Campan M, Long TI *et al.* Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Res* 2005;**33**:6823-36.

[9] Yang AS, Estecio MR, Doshi K, Kondo Y, Tajara EH, Issa JP. A simple method for estimating global DNA methylation using bisulfite PCR of repetitive DNA elements. *Nucleic Acids Res* 2004;**32**:e38.

[10] Baccarelli A, Wright RO, Bollati V *et al.* Rapid DNA methylation changes after exposure to traffic particles. *Am J Respir Crit Care Med* 2009;**179**:572-8.

[11] Baccarelli A, Wright RO, Bollati V *et al.* Low blood DNA methylation determines higher risk and mortality from ischemic heart disease and stroke among elderly individuals. *Circulation* 2009;**119**:E281.


For Review Only

[12] Bollati V, Baccarelli A, Hou L *et al.* Changes in DNA methylation patterns in subjects exposed to low-dose benzene. *Cancer Res* 2007;**67**:876-80.

[13] Rusiecki JA, Baccarelli A, Bollati V, Tarantini L, Moore LE, Bonefeld-Jorgensen EC. Global DNA hypomethylation is associated with high serum-persistent organic pollutants in Greenlandic Inuit. *Environ Health Perspect* 2008;**116**:1547-52.

[14] Bollati V, Schwartz J, Wright R *et al.* Decline in genomic DNA methylation through aging in a cohort of elderly subjects. *Mech Ageing Dev* 2009;**130**:234-9.

[15] El-Maarri O, Becker T, Junen J *et al.* Gender specific differences in levels of DNA methylation at selected loci from human total blood: a tendency toward higher methylation levels in males. *Hum Genet* 2007;**122**:505-14.

[16] Ogino S, Nosho K, Kirkner GJ *et al.* A cohort study of tumoral LINE-1 hypomethylation and prognosis in colon cancer. *J Natl Cancer Inst* 2008;**100**:1734-8.

[17] Tarantini L, Bonzini M, Apostoli P *et al.* Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation. *Environ Health Perspect* 2009;**117**:217-22.

[18] Figueiredo JC, Grau MV, Wallace K *et al.* Global DNA hypomethylation (LINE-1) in the normal colon and lifestyle characteristics and dietary and genetic factors. *Cancer Epidemiol Biomarkers Prev* 2009;**18**:1041-9.

[19] Smith IM, Mydlarz WK, Mithani SK, Califano JA. DNA global hypomethylation in squamous cell head and neck cancer associated with smoking, alcohol consumption and stage. *Int J Cancer* 2007;**121**:1724-8.

[20] Richards KL, Zhang B, Baggerly KA *et al.* Genome-wide hypomethylation in head and neck cancer is more pronounced in HPV-negative tumors and is associated with genomic instability. *PLoS One* 2009;**4**:e4941.

For Review Only

[21] Yamamoto E, Toyota M, Suzuki H *et al.* LINE-1 hypomethylation is associated with increased CpG island methylation in Helicobacter pylori-related enlarged-fold gastritis. *Cancer Epidemiol Biomarkers Prev* 2008;**17**:2555-64.

[22] Woloszynska-Read A, Mhawech-Fauceglia P, Yu J, Odunsi K, Karpf AR. Intertumor and intratumor NY-ESO-1 expression heterogeneity is associated with promoter-specific and global DNA methylation status in ovarian cancer. *Clin Cancer Res* 2008;**14**:3283-90.

[23] Geli J, Kiss N, Karimi M *et al.* Global and regional CpG methylation in pheochromocytomas and abdominal paragangliomas: association to malignant behavior. *Clin Cancer Res* 2008;**14**:2551-9.

[24] Ogino S, Kawasaki T, Nosho K *et al.* LINE-1 hypomethylation is inversely associated with microsatellite instability and CpG island methylator phenotype in colorectal cancer. *Int J Cancer* 2008;**122**:2767-73.

[25] Iacopetta B, Grieu F, Phillips M *et al.* Methylation levels of LINE-1 repeats and CpG island loci are inversely related in normal colonic mucosa. *Cancer Sci* 2007;**98**:1454-60.

[26] Choi IS, Estecio MR, Nagano Y *et al.* Hypomethylation of LINE-1 and Alu in well-differentiated neuroendocrine tumors (pancreatic endocrine tumors and carcinoid tumors). *Mod Pathol* 2007;**20**:802-10.

[27] Estecio MR, Gharibyan V, Shen L *et al.* LINE-1 hypomethylation in cancer is highly variable and inversely correlated with microsatellite instability. *PLoS One* 2007;**2**:e399.

[28] Lucchinetti E, Feng J, Silva R *et al.* Inhibition of LINE-1 expression in the heart decreases ischemic damage by activation of Akt/PKB signaling. *Physiol Genomics* 2006;**25**:314-24.

[29] Castro R, Rivera I, Struys EA *et al.* Increased homocysteine and S-adenosylhomocysteine concentrations and DNA hypomethylation in vascular disease. *Clin Chem* 2003;**49**:1292-6.

[30] Wilson AS, Power BE, Molloy PL. DNA hypomethylation and human diseases. *Biochim Biophys Acta* 2007;**1775**:138-62.

For Review Only

[31] Foley DL, Craig JM, Morley R *et al.* Prospects for epigenetic epidemiology. *Am J Epidemiol* 2009;**169**:389-400.

[32] Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003;**33**:245-54.

[33] Bell B, Rose C, Damon A. The normative aging study: an interdisciplinary and longitudinal study of health and aging. *Aging Hum Dev* 1972;**3**:4–17.

[34] Chow WH, Swanson CA, Lissowska J *et al.* Risk of stomach cancer in relation to consumption of cigarettes, alcohol, tea and coffee in Warsaw, Poland. *Int J Cancer* 1999;**81**:871-6.

[35] Costa G, Sartori S, Consonni D. Thirty years of medical surveillance in perfluooctanoic acid production workers. *J Occup Environ Med* 2009;**51**:364-72.

[36] McCleary R, Dick MB, Buckwalter G, Henderson V, Shankle WR. Full-information models for multiple psychometric tests: annualized rates of change in normal aging and dementia. *Alzheimer Dis Assoc Disord* 1996;**10**:216-23.

[37] Li TH, Schmid CW. Differential stress induction of individual Alu loci: implications for transcription and retrotransposition. *Gene* 2001;**276**:135-41.

[38] Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 2002;**3**:415-28.

[39] Fraga MF, Agrelo R, Esteller M. Cross-talk between aging and cancer: the epigenetic language. *Ann N Y Acad Sci* 2007;**1100**:60-74.

[40] Liu L, Wylie RC, Andrews LG, Tollefsbol TO. Aging, cancer and nutrition: the DNA methylation connection. *Mech Ageing Dev* 2003;**124**:989-98.

[41] Fuke C, Shimabukuro M, Petronis A *et al.* Age related changes in 5-methylcytosine content in human peripheral leukocytes and placentas: an HPLC-based study. *Ann Hum Genet* 2004;**68**:196-204.

For Review Only

[42] Shimabukuro M, Sasaki T, Imamura A *et al.* Global hypomethylation of peripheral leukocyte DNA in male patients with schizophrenia: a potential link between epigenetics and schizophrenia. *J Psychiatr Res* 2007;**41**:1042-6.

[43] Moore LE, Pfeiffer RM, Poscablo C *et al.* Genomic DNA hypomethylation as a biomarker for bladder cancer susceptibility in the Spanish Bladder Cancer Study: a case-control study. *Lancet Oncol* 2008;**9**:359-66.

[44] Chalitchagorn K, Shuangshoti S, Hourpai N *et al.* Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. *Oncogene* 2004;**23**:8841-6.

[45] Hsiung DT, Marsit CJ, Houseman EA *et al.* Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. *Cancer Epidemiol Biomarkers Prev* 2007;**16**:108-14.

[46] Bjornsson HT, Sigurdsson MI, Fallin MD *et al.* Intra-individual change over time in DNA methylation with familial clustering. *JAMA* 2008;**299**:2877-83.

[47] Dunn BK. Hypomethylation: one side of a larger picture. *Ann N Y Acad Sci* 2003;**983**:28-42.

[48] Furniss CS, Marsit CJ, Houseman EA, Eddy K, Kelsey KT. Line region hypomethylation is associated with lifestyle and differs by human papillomavirus status in head and neck squamous cell carcinomas. *Cancer Epidemiol Biomarkers Prev* 2008;**17**:966-71.

[49] Guerrero-Preston R, Baez A, Blanco A, Berdasco M, Fraga M, Esteller M. Global DNA methylation: a common early event in oral cancer cases with exposure to environmental carcinogens or viral agents. *P R Health Sci J* 2009;**28**:24-9.

[50] Vaissiere T, Hung RJ, Zaridze D *et al.* Quantitative analysis of DNA methylation profiles in lung cancer identifies aberrant DNA methylation of specific genes and its association with gender and cancer risk factors. *Cancer Res* 2009;**69**:243-52.

For Review Only

[51] Hillemacher T, Frieling H, Moskau S *et al.* Global DNA methylation is influenced by smoking behaviour. *Eur Neuropsychopharmacol* 2008;**18**:295-8.

[52] Breton CV, Byun HM, Wenten M, Pan F, Yang A, Gilliland FD. Prenatal Tobacco Smoke Exposure Affects Global and Gene-Specific DNA Methylation. *Am J Respir Crit Care Med* 2009;**180**:462-7.

[53] Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions on adult health and disease. *N Engl J Med* 2008;**359**:61-73.

[54] Gibson A, Woodside JV, Young IS *et al.* Alcohol increases homocysteine and reduces B vitamin concentration in healthy male volunteers--a randomized, crossover intervention study. *QJM* 2008;**101**:881-7.

[55] Purohit V, Abdelmalek MF, Barve S *et al.* Role of S-adenosylmethionine, folate, and betaine in the treatment of alcoholic liver disease: summary of a symposium. *Am J Clin Nutr* 2007;**86**:14-24.

[56] Bielawski DM, Zaher FM, Svinarich DM, Abel EL. Paternal alcohol exposure affects sperm cytosine methyltransferase messenger RNA levels. *Alcohol Clin Exp Res* 2002;**26**:347-51.

[57] Fiala ES, Sohn OS, Hamilton SR. Effects of chronic dietary ethanol on in vivo and in vitro metabolism of methylazoxymethanol and on methylazoxymethanol-induced DNA methylation in rat colon and liver. *Cancer Res* 1987;**47**:5939-43.

[58] Eckhardt F, Lewin J, Cortese R *et al.* DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* 2006;**38**:1378-85.

[59] Von Essen S, Romberger D. The respiratory inflammatory response to the swine confinement building environment: the adaptation to respiratory exposures in the chronically exposed worker. *J Agric Saf Health* 2003;**9**:185-96.

[60] Requena L, Sanchez Yus E. Erythema nodosum. *Semin Cutan Med Surg* 2007;**26**:114-25.

For Review Only

[61] Piyathilake CJ, Henao O, Frost AR *et al.* Race- and age-dependent alterations in global methylation of DNA in squamous cell carcinoma of the lung (United States). *Cancer Causes Control* 2003;**14**:37-42.

[62] Suzuki H, Toyota M, Sato H, Sonoda T, Sakauchi F, Mori M. Roles and causes of abnormal DNA methylation in gastrointestinal cancers. *Asian Pac J Cancer Prev* 2006;**7**:177-85.

For Review Only

**Table 1** Subject characteristics of the five studies included in a combined analysis of determinants of blood DNA methylation in repetitive elements in healthy individuals [mean or *n* (%)]

| Variable | Study #1 Boston, USA | Study #2 Warsaw, Poland | Study #3 Milan, Italy | Study #4 Brescia, Italy | Study #5 Trissino, Italy | All studies (*n*=1465) |
|---|---|---|---|---|---|---|
| **Age, years (min-max)** | 72 (55-100) | 63 (28-80) | 37 (19-74) | 44 (27-55) | 43 (26-62) | 62 (19-100) |
| **Gender, *n* (%)** | | | | | | |
| Female | 0 (0.0) | 149 (35.4) | 58 (27.5) | 0 (0.0) | 0 (0.0) | 207 (14.1) |
| Male | 722 (100.0) | 272 (64.6) | 153 (72.5) | 63 (100.0) | 48 (100.0) | 1258 (85.9) |
| **Body mass index, kg/m$^2$ (SD)** | 28.2 (4.1) | 25.7 (3.7) | N/A | 26.5 (2.7) | 25.5 (2.9) | 27.2 (4.0) |
| **Smoking status, *n* (%)** | | | | | | |
| Nonsmoker | 212 (29.4) | 169 (40.1) | 95 (45.0) | 20 (33.3) | 28 (58.3) | 524 (35.8) |
| Former smoker | 482 (66.8) | 143 (34.0) | 33 (15.6) | 15 (25.0) | 2 (4.2) | 675 (46.2) |
| Current smoker | 28 (3.9) | 109 (25.9) | 83 (39.3) | 25 (41.7) | 18 (37.5) | 263 (18.0) |
| **Current cigarettes, *n*/day (min-max)** | 20 (3-40) | 18 (2-97) | 15 (1-40) | 13 (1-30) | N/A | 17 (1-97) |
| **Past cigarettes, *n*/day (min-max)** | 26 (1-144) | 21 (2-97) | 19 (2-60) | 17 (5-50) | N/A | 24 (1-144) |
| **Total smoking years (min-max)** | 24 (1-63) | 30 (1-59) | 19 (1-48) | 21 (3-38) | N/A | 25 (1-63) |
| **Pack-years of smoking (min-max)** | 32 (0.1-267) | 31 (0.1-162) | 17 (0.3-82) | 16 (0.7-50) | N/A | 29 (0.1-267) |
| **Age smoking started, years (min-max)** | 18 (6-44) | 20 (11-73) | 17 (13-28) | 18 (10-30) | N/A | 18 (6-73) |
| **Age smoking stopped, years (min-max)** | 42 (13-84) | 49 (19-75) | 33 (19-56) | 34 (20-53) | N/A | 43 (13-84) |
| **Alcohol drinking, *n* (%)** | | | | | | |
| Nondrinker | 390 (58.4) | 138 (32.8) | 77 (36.8) | N/A | 14 (29.2) | 619 (46.0) |
| Ever-drinker | 278 (41.6) | 283 (67.2) | 132 (63.2) | N/A | 34 (70.8) | 727 (54.0) |
| **Blood count** | | | | | | |
| White blood cells, 10$^3$cell/mm$^3$(min-max) | 6.6 (1.8-16.4) | N/A | 7.1 (3.5-18.5) | 7.3 (3.6-10.4) | 7.0 (4.1-12.6) | 6.6 (1.8-18.5) |
| Neutrophils, % (min-max) | 62.2 (31-85) | N/A | 54.5 (33.7-84.3) | N/A | 53.3 (34.7-87.9) | 60.1 (31-87.9) |
| Lymphocytes, % (min-max) | 25.5(5-65) | N/A | 34.3 (8.5-55.5) | 30.7 (15.4-44.1) | 36.1 (7.7-60.1) | 28.2 (5-65) |
| Monocytes, % (min-max) | 8.7 (2-17) | N/A | 8.0 (2.6-14.5) | 10.4 (4.5-19.4) | 9.1 (4.2-18.7) | 8.7 (2-19.4) |
| Eosinophils, % (min-max) | 3.3 (0-22) | N/A | 2.9 (0.1-9.2) | N/A | 2.8 (0.1-7.1) | 3.2 (0-22) |
| Basophils, % (min-max) | 0.6 (0-2) | N/A | 0.3 (0-1.5) | N/A | 0.4 (0-1.4) | 0.5 (0-2) |
| **DNA methylation, %5mC (SD)** | | | | | | |
| Alu | 26.3 (1.1) | 25.1 (1.4) | 27.1 (3.4) | 25.8 (0.7) | 24.9 (1.0) | 26.0 (1.8) |
| LINE-1 | 76.9 (2.0) | 80.2 (3.2) | 71.7 (5.9) | 78.8 (1.0) | 78.7 (3.2) | 76.2 (6.2) |

Abbreviations: %5mC, percentage of 5-methyl cytosine; max, maximum; min, minimum; N/A, not available; SD, standard deviation.

For Review Only

**Table 2** Association of Alu and LINE-1 methylation with the characteristics of the study subjects, in combination as well as within each of the five studies

| | Variable | Alu | | | | LINE-1 | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | No. | %5mC[a] | 95% CI[a] | Beta (95% CI)[a] | No. | %5mC[a] | 95% CI[a] | Beta (95% CI)[a] |
| **All studies** | **Age, years** | | | | | | | | |
| | <48 | 285 | 26.1 | 25.8, 26.4 | | 264 | 77.2 | 76.7, 77.8 | |
| | 48-63 | 279 | 26.0 | 25.8, 26.2 | | 271 | 77.4 | 77.0, 77.8 | |
| | 64-69 | 314 | 26.0 | 25.8, 26.2 | | 307 | 77.2 | 76.8, 77.6 | |
| | 70-74 | 267 | 25.9 | 25.7, 26.1 | | 263 | 77.1 | 76.7, 77.6 | |
| | ≥75 | 320 | 25.9 | 25.7, 26.1 | −0.011 (−0.020 to −0.001)[b] | 313 | 77.2 | 76.8, 77.6 | 0.003 (−0.016 to 0.021)[b] |
| | **Gender** | | | | | | | | |
| | Female | 206 | 26.3 | 26.1, 26.6 | | 202 | 76.6 | 76.1, 77.0 | |
| | Male | 1238 | 25.9 | 25.8, 26.0 | −0.385 (−0.665 to −0.104) | 1216 | 77.4 | 77.2, 77.5 | 0.796 (0.261 to 1.330) |
| | **Body mass index, kg/m$^2$** | | | | | | | | |
| | <25.3 | 413 | 25.8 | 25.7, 25.9 | | 408 | 78.0 | 77.8, 78.3 | |
| | 25.3-28.5 | 415 | 25.8 | 25.6, 25.9 | | 415 | 78.3 | 78.1, 78.5 | |
| | ≥28.6 | 410 | 25.9 | 25.7, 26.0 | 0.012 (−0.005 to 0.030)[b] | 408 | 78.2 | 78.0, 78.4 | 0.003 (−0.033 to 0.040)[b] |
| | **Alcohol drinking** | | | | | | | | |
| | Nondrinker | 609 | 26.1 | 26.0, 26.3 | | 604 | 77.2 | 76.9, 77.4 | |
| | Ever-drinker | 717 | 25.9 | 25.8, 26.0 | −0.214 (−0.415 to −0.013) | 700 | 77.1 | 76.9, 77.3 | −0.053 (−0.430 to 0.323) |
| | **Smoking status** | | | | | | | | |
| | Nonsmoker | 519 | 26.0 | 25.8, 26.1 | | 504 | 77.2 | 77.0, 77.5 | |
| | Former smoker | 665 | 26.0 | 25.9, 26.1 | | 659 | 77.3 | 77.0, 77.5 | |
| | Current smoker | 257 | 26.0 | 25.8, 26.2 | | 252 | 77.2 | 76.7, 77.6 | |
| | Trend statistics | | | | 0.018 (−0.107 to 0.143) | | | | −0.025 (−0.263 to 0.214) |
| **Study #1** | **Age, years** | | | | | | | | |
| | <48 | 0 | NA | NA | | 0 | NA | NA | |
| | 48-63 | 63 | 26.5 | 26.2, 26.7 | | 63 | 77.4 | 77.0, 77.9 | |
| | 64-69 | 199 | 26.5 | 26.3, 26.6 | | 196 | 76.9 | 76.6, 77.1 | |
| | 70-74 | 185 | 26.3 | 26.1, 26.4 | | 186 | 76.8 | 76.5, 77.1 | |
| | ≥75 | 261 | 26.2 | 26.1, 26.3 | −0.022 (−0.034 to −0.010)[b] | 258 | 76.9 | 76.6, 77.1 | −0.011 (−0.033 to 0.010)[b] |
| | **Gender** | | | | | | | | |
| | Female | 0 | NA | NA | | 0 | NA | NA | |



For Review Only

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Male | 708 | 26.3 | 26.2, 26.4 | NA | 703 | 76.9 | 76.8, 77.0 | NA |
| | **Body mass index, kg/m$^2$** | | | | | | | | |
| | <25.3 | 160 | 26.3 | 26.1, 26.4 | | 156 | 76.9 | 76.6, 77.2 | |
| | 25.3-28.5 | 251 | 26.3 | 26.1, 26.4 | | 252 | 77.0 | 76.7, 77.2 | |
| | ≥28.6 | 297 | 26.4 | 26.2, 26.5 | 0.021 (0.001 to 0.041)[b] | 295 | 76.9 | 76.6, 77.1 | 0.005 (−0.032 to 0.042)[b] |
| | **Alcohol drinking** | | | | | | | | |
| | Nondrinker | 383 | 26.3 | 26.2, 26.5 | | 381 | 76.7 | 76.5, 76.9 | |
| | Ever-drinker | 271 | 26.3 | 26.2, 26.4 | −0.023 (−0.188 to 0.143) | 272 | 76.9 | 76.7, 77.1 | 0.239 (−0.037 to 0.515) |
| | **Smoking status** | | | | | | | | |
| | Nonsmoker | 208 | 26.3 | 26.2, 26.5 | | 206 | 76.7 | 76.5, 77.0 | |
| | Former smoker | 473 | 26.3 | 26.2, 26.4 | | 470 | 76.9 | 76.8, 77.1 | |
| | Current smoker | 27 | 26.2 | 25.8, 26.6 | | 27 | 77.4 | 76.7, 78.2 | |
| | Trend statistics | | | | −0.015 (−0.173 to 0.143) | | | | 0.252 (−0.033 to 0.536) |
| **Study #2** | **Age, years** | | | | | | | | |
| | <48 | 41 | 25.0 | 24.6, 25.5 | | 40 | 79.4 | 78.4, 80.4 | |
| | 48-63 | 141 | 25.1 | 24.8, 25.3 | | 141 | 80.5 | 79.9, 81.0 | |
| | 64-69 | 109 | 24.9 | 24.7, 25.2 | | 108 | 80.2 | 79.6, 80.9 | |
| | 70-74 | 75 | 24.9 | 24.6, 25.3 | | 75 | 80.2 | 79.5, 80.9 | |
| | ≥75 | 55 | 25.4 | 25.0, 25.8 | 0.004 (−0.009 to 0.017)[b] | 55 | 80.2 | 79.4, 81.1 | 0.017 (−0.012 to 0.047)[b] |
| | **Gender** | | | | | | | | |
| | Female | 149 | 25.2 | 25.0, 25.4 | | 148 | 79.6 | 79.1, 80.1 | |
| | Male | 272 | 25.0 | 24.8, 25.1 | −0.201 (−0.486 to 0.084) | 271 | 80.6 | 80.2, 80.9 | 0.951 (0.304 to 1.597) |
| | **Body mass index, kg/m$^2$** | | | | | | | | |
| | <25.3 | 206 | 25.1 | 24.9, 25.3 | | 205 | 80.1 | 79.6, 80.5 | |
| | 25.3-28.5 | 126 | 25.0 | 24.7, 25.2 | | 125 | 80.4 | 79.8, 81.0 | |
| | ≥28.6 | 88 | 25.0 | 24.7, 25.3 | −0.013 (−0.051 to 0.025)[b] | 88 | 80.3 | 79.6, 81.0 | −0.024 (−0.110 to 0.062)[b] |
| | **Alcohol drinking** | | | | | | | | |
| | Nondrinker | 138 | 25.2 | 24.9, 25.5 | | 138 | 80.4 | 79.8, 81.0 | |
| | Ever-drinker | 283 | 25.0 | 24.8, 25.2 | −0.227 (−0.567 to 0.112) | 281 | 80.2 | 79.8, 80.6 | −0.218 (−0.991 to 0.554) |
| | **Smoking status** | | | | | | | | |
| | Nonsmoker | 169 | 24.9 | 24.7, 25.1 | | 168 | 80.3 | 79.7, 80.8 | |
| | Former smoker | 143 | 25.1 | 24.9, 25.4 | | 143 | 80.3 | 79.7, 80.8 | |
| | Current smoker | 109 | 25.2 | 24.9, 25.5 | | 108 | 80.1 | 79.5, 80.8 | |

For Review Only

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Trend statistics | | | | 0.136 (−0.046 to 0.317) | | | | −0.053 (−0.466 to 0.361) |
| **Study #3** | **Age, years** | | | | | | | | |
| | <48 | 164 | 27.2 | 26.7, 27.7 | | 150 | 71.8 | 70.8, 72.7 | |
| | 48-63 | 36 | 26.7 | 25.6, 27.8 | | 31 | 71.3 | 69.2, 73.4 | |
| | 64-69 | 3 | 26.8 | 23.0, 30.7 | | 3 | 72.7 | 65.9, 79.4 | |
| | 70-74 | 2 | 27.4 | 22.7, 32.1 | | 2 | 69.5 | 61.1, 77.8 | |
| | ≥75 | 0 | NA | NA | −0.027 (−0.069 to 0.014)[b] | 0 | NA | NA | 0.007 (−0.071 to 0.084)[b] |
| | **Gender** | | | | | | | | |
| | Female | 57 | 27.6 | 26.8, 28.5 | | 54 | 71.4 | 69.8, 73.0 | |
| | Male | 148 | 26.9 | 26.3, 27.4 | −0.782 (−1.811 to 0.247) | 132 | 71.8 | 70.8, 72.8 | 0.397 (−1.500 to 2.294) |
| | **Body mass index, kg/m$^2$** | | | | | | | | |
| | <25.3 | NA | NA | NA | | NA | NA | NA | |
| | 25.3-28.5 | NA | NA | NA | | NA | NA | NA | |
| | ≥28.6 | NA | NA | NA | NA | NA | NA | NA | NA |
| | **Alcohol drinking** | | | | | | | | |
| | Nondrinker | 74 | 27.6 | 26.7, 28.4 | | 71 | 72.3 | 70.8, 73.8 | |
| | Ever-drinker | 129 | 26.8 | 26.2, 27.4 | −0.756 (−1.831 to 0.319) | 113 | 71.2 | 70.0, 72.3 | −1.164 (−3.128 to 0.800) |
| | **Smoking status** | | | | | | | | |
| | Nonsmoker | 94 | 27.2 | 26.5, 27.8 | | 82 | 72.0 | 70.7, 73.3 | |
| | Former smoker | 32 | 26.9 | 25.7, 28.1 | | 29 | 71.2 | 69.0, 73.4 | |
| | Current smoker | 79 | 27.1 | 26.3, 27.8 | | 75 | 71.5 | 70.2, 72.9 | |
| | Trend statistics | | | | −0.051 (−0.554 to 0.451) | | | | −0.232 (−1.171 to 0.707) |
| **Study #4** | **Age, years** | | | | | | | | |
| | <48 | 43 | 25.8 | 25.5, 26.0 | | 43 | 78.8 | 78.5, 79.1 | |
| | 48-63 | 19 | 26.0 | 25.7, 26.3 | | 19 | 78.9 | 78.4, 79.3 | |
| | 64-69 | 0 | NA | NA | | 0 | NA | NA | |
| | 70-74 | 0 | NA | NA | | 0 | NA | NA | |
| | ≥75 | 0 | NA | NA | 0.014 (−0.010 to 0.039)[b] | 0 | NA | NA | −0.006 (−0.040 to 0.027)[b] |
| | **Gender** | | | | | | | | |
| | Female | 0 | NA | NA | | 0 | NA | NA | |
| | Male | 62 | 25.8 | 25.7, 26.0 | NA | 62 | 78.8 | 78.6, 79.1 | NA |
| | **Body mass index, kg/m$^2$** | | | | | | | | |
| | <25.3 | 21 | 25.7 | 25.4, 26.0 | | 21 | 79.1 | 78.7, 79.5 | |

For Review Only

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 25.3-28.5 | 26 | 25.8 | 25.5, 26.1 | | 26 | 78.8 | 78.4, 79.2 | |
| | ≥28.6 | 15 | 26.0 | 25.6, 26.4 | 0.035 (−0.034 to 0.104)[b] | 15 | 78.4 | 77.9, 78.9 | −0.108 (−0.201 to −0.016)[b] |
| | **Alcohol drinking** | | | | | | | | |
| | Nondrinker | NA | NA | NA | | NA | NA | NA | |
| | Ever-drinker | NA | NA | NA | NA | NA | NA | NA | NA |
| | **Smoking status** | | | | | | | | |
| | Nonsmoker | 20 | 26.2 | 25.9, 26.5 | | 20 | 78.8 | 78.3, 79.2 | |
| | Former smoker | 15 | 25.5 | 25.1, 25.8 | | 15 | 78.8 | 78.3, 79.4 | |
| | Current smoker | 24 | 25.9 | 25.6, 26.1 | | 24 | 78.9 | 78.4, 79.3 | |
| | Trend statistics | | | | −0.162 (−0.380 to 0.055) | | | | 0.057 (−0.259 to 0.373) |
| **Study #5** | **Age, years** | | | | | | | | |
| | <48 | 31 | 24.9 | 24.6, 25.3 | | 31 | 79.1 | 77.9, 80.2 | |
| | 48-63 | 17 | 24.8 | 24.3, 25.3 | | 17 | 78.1 | 76.6, 79.7 | |
| | 64-69 | 0 | NA | NA | | 0 | NA | NA | |
| | 70-74 | 0 | NA | NA | | 0 | NA | NA | |
| | ≥75 | 0 | NA | NA | 0.006 (−0.025 to 0.037)[b] | 0 | NA | NA | −0.061 (−0.161 to 0.039)[b] |
| | **Gender** | | | | | | | | |
| | Female | 0 | NA | NA | | 0 | NA | NA | |
| | Male | 48 | 24.9 | 24.6, 25.2 | NA | 48 | 78.7 | 77.8, 79.7 | NA |
| | **Body mass index, kg/m$^2$** | | | | | | | | |
| | <25.3 | 26 | 25.0 | 24.6, 25.4 | | 26 | 77.6 | 76.4, 78.8 | |
| | 25.3-28.5 | 12 | 24.6 | 24.1, 25.2 | | 12 | 79.7 | 77.9, 81.5 | |
| | ≥28.6 | 10 | 24.8 | 24.1, 25.5 | −0.093 (−0.198 to 0.011)[b] | 10 | 80.5 | 78.5, 82.5 | 0.386 (0.058 to 0.714)[b] |
| | **Alcohol drinking** | | | | | | | | |
| | Nondrinker | 14 | 25.4 | 24.8, 25.9 | | 14 | 77.6 | 75.8, 79.5 | |
| | Ever-drinker | 34 | 24.7 | 24.3, 25.0 | −0.687 (−1.379 to 0.005) | 34 | 79.2 | 78.0, 80.3 | 1.557 (−0.722 to 3.836) |
| | **Smoking status** | | | | | | | | |
| | Nonsmoker | 28 | 24.9 | 24.5, 25.3 | | 28 | 79.1 | 77.9, 80.3 | |
| | Former smoker | 2 | 25.3 | 23.8, 26.8 | | 2 | 74.5 | 70.0, 79.1 | |
| | Current smoker | 18 | 24.8 | 24.3, 25.3 | | 18 | 78.6 | 77.1, 80.2 | |
| | Trend statistics | | | | −0.027 (−0.341 to 0.287) | | | | −0.290 (−1.300 to 0.719) |

Abbreviations: %5mC, percentage of 5-methyl cytosine; NA, not available.
[a] Adjusted for age, gender and study.
[b] Beta for age and body mass index were obtained using continuous variables.

For Review Only

**Table 3** Effect of blood count on Alu and LINE-1 methylation (combined analysis of the five studies)

| Blood count | Alu | | LINE-1 | |
|---|---|---|---|---|
| | No. | Beta (95% CI) [a] | No. | Beta (95% CI) [a] |
| **White blood cells, $10^3$cell/mm$^3$** | 1013 | 0.001 (−0.060 to 0.062) | 991 | 0.077 (−0.034 to 0.187) |
| **Neutrophils, %** | 939 | 0.008 (−0.006 to 0.022) | 917 | 0.036 (0.010 to 0.061) |
| **Lymphocytes, %** | 1001 | −0.009 (−0.023 to 0.006) | 979 | −0.038 (−0.065 to −0.012) |
| **Monocytes, %** | 994 | 0.001 (−0.040 to 0.042) | 972 | −0.032 (−0.105 to 0.042) |
| **Eosinophils, %** | 932 | −0.012 (−0.070 to 0.047) | 910 | 0.010 (−0.094 to 0.114) |
| **Basophils, %** | 932 | 0.008 (−0.254 to 0.270) | 910 | −0.198 (−0.668 to 0.272) |

[a] Adjusted for age, gender, alcohol drinking and study.

For Review Only

**Table 4** Effects of age, gender and alcohol drinking on Alu and LINE-1 methylation in multivariable models including percent lymphocytes in blood count, age, gender and study as independent variables (combined analysis of the five studies)

| Variable | Alu | | | | LINE-1 | | | |
|---|---|---|---|---|---|---|---|---|
| | No. | %5mC[a] | 95% CI[a] | Beta (95% CI)[a] | No. | %5mC[a] | 95% CI[a] | Beta (95% CI)[a] |
| **Age, years** | | | | | | | | |
| <48 | 285 | 26.2 | 25.9, 26.5 | | 264 | 77.4 | 76.8, 77.9 | |
| 48-63 | 279 | 26.0 | 25.8, 26.2 | | 271 | 77.5 | 77.1, 77.9 | |
| 64-69 | 314 | 26.0 | 25.8, 26.2 | | 307 | 77.2 | 76.8, 77.6 | |
| 70-74 | 267 | 25.9 | 25.7, 26.1 | | 263 | 77.1 | 76.7, 77.5 | |
| ≥75 | 320 | 25.9 | 25.7, 26.1 | −0.012 (−0.022 to −0.002)[b] | 313 | 77.1 | 76.7, 77.5 | −0.003 (−0.022 to 0.016)[b] |
| **Gender** | | | | | | | | |
| Female | 206 | 26.3 | 26.0, 26.6 | | 202 | 76.5 | 76.0, 77.0 | |
| Male | 1238 | 25.9 | 25.8, 26.0 | −0.361 (−0.642 to −0.079) | 1216 | 77.4 | 77.2, 77.5 | 0.874 (0.339 to 1.409) |
| **Alcohol drinking** | | | | | | | | |
| Nondrinker | 609 | 26.1 | 26.0, 26.3 | | 604 | 77.2 | 76.9, 77.4 | |
| Ever-drinker | 717 | 25.9 | 25.8, 26.0 | −0.208 (−0.409 to −0.007) | 700 | 77.1 | 76.9, 77.3 | −0.049 (−0.425 to 0.327) |

Abbreviations: %5mC, percentage of 5-methyl cytosine.
[a] Adjusted for age, gender, percent lymphocytes in blood count, and study.
[b] Beta for age were obtained using continuous variables.

For Review Only