Particulate Air Pollution and Fasting Blood Glucose in Nondiabetic Individuals: Associations and Epigenetic Mediation in the Normative Aging Study, 2000–2011

Cheng Peng,1 Marie-Abele C. Bind,2 Elena Colicino,1 Itai Kloo,3 Hyang-Min Byun,4 Laura Cantone,5 Letizia Trevisi,1 Jia Zhong,1 Kasey Brennan,1 Alexandra E. Dereix,1 Pantel S. Vokonas,6,7 Brent A. Coull,8 Joel D. Schwartz,1,9 and Andrea A. Baccarelli1

1Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA; 2Department of Environmental, University of Cambridge, Cambridge, Massachusetts, USA; 3Department of Geography and Environmental Development, Ben-Gurion University of the Negev, Beer Sheba, Israel; 4Human Nutrition Research Center, Institute of Cellular Medicine, Newcastle University, Newcastle, United Kingdom; 5Molecular Epidemiology and Environmental Epigenetics Laboratory, Department of Clinical Sciences and Community Health, University of Milan, Milan, Italy; 6VA Normative Aging Study, Veterans Affairs Boston Healthcare System, Boston, Massachusetts, USA; 7Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA; 8Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA; 9Channing Laboratory, Harvard Medical School, Boston, Massachusetts, USA

BACKGROUND: Among nondiabetic individuals, higher fasting blood glucose (FBG) independently predicts diabetes risk, cardiovascular disease, and dementia. Ambient PM2.5 (particulate matter with aerodynamic diameter ≤ 2.5 μm) is an emerging determinant of glucose dysregulation. PM2.5 effects and mechanisms are understudied among nondiabetic individuals.

OBJECTIVES: Our goals were to investigate whether PM2.5 is associated with an increase in FBG and to explore potential mediating roles of epigenetic gene regulation.

METHODS: In 551 nondiabetic participants in the Normative Aging Study, we measured FBG, and DNA methylation of four inflammatory genes (IL-1β, IL-6, ICAM-1, and TLR-2), up to four times between 2000 and 2011 (median = 2). We estimated short- and medium-term (1-, 7-, and 28-day preceding each clinical visit) ambient PM2.5 at each participant’s address using a validated hybrid land-use regression satellite-based model. We fitted covariate-adjusted regression models accounting for repeated measures.

RESULTS: Mean FBG was 99.8 mg/dL (SD = 10.7), 18% of the participants had impaired fasting glucose (IFG; i.e., 100–125 mg/dL FBG) at first visit. Interquartile increases in 1-, 7-, and 28-day PM2.5 were associated with 0.57 mg/dL (95% CI: 0.02, 1.11, p = 0.04), 1.02 mg/dL (95% CI: 0.41, 1.63, p = 0.001), and 0.89 mg/dL (95% CI: 0.32, 1.47, p = 0.003) higher FBG, respectively. The same PM2.5 metrics were associated with 13% (95% CI: −3%, 33%, p = 0.12), 27% (95% CI: 6%, 52%, p = 0.01) and 32% (95% CI: 10%, 58%, p = 0.003) higher odds of IFG, respectively. PM2.5 was negatively correlated with ICAM-1 methylation (p = 0.01), but not with other genes. Mediation analysis estimated that ICAM-1 methylation mediated 9% of the association of 28-day PM2.5 with FBG.

CONCLUSIONS: Among nondiabetics, short- and medium-term PM2.5 were associated with higher FBG. Mediation analysis indicated that part of this association was mediated by ICAM-1 promoter methylation.


Introduction

Clinically diagnosed diabetes is preceded by a long latent period of abnormal glucose metabolism (American Diabetes Association 2012). In the asymptomatic, nondiabetic range of glycemia (<126 mg/dL) (American Diabetes Association 2012), increased fasting blood glucose (FBG) levels are already independently associated with the development of diabetes (Tirosh et al. 2005), cardiovascular disease (Duckworth 2001; Levitan et al. 2004; Sarwar et al. 2010), and dementia (Crane et al. 2013). FBG variations in this range are often underestimated and their potential determinants, especially those not directly related to lifestyle, are understudied. Ambient particulate matter (PM) pollution has recently been suggested as an emerging risk factor for metabolic disorders including impaired glucose regulation (Esposito et al. 2016; Rajagopalan and Brook 2012). Cross-sectional and longitudinal studies have revealed that PM is associated with increased risk of diabetes (Andersen et al. 2012; Brook et al. 2008; Puett et al. 2011) and with higher levels of markers of insulin resistance (Kelishadi et al. 2009). However, the relationship between PM and glycemia in the nondiabetic range has yet to be studied.

The underlying mechanisms linking PM and abnormal glucose regulation are also not fully understood. Inflammation is central in both PM-associated responses and the pathogenesis of glucose dysregulation. Evidence from previous studies has linked exposures to ambient PM with lower DNA methylation in inflammatory genes (Bind et al. 2012, 2014). DNA methylation, through the addition of a methyl group to the 5′ position of cytosine in the CpG dinucleotide sequence, is a well-studied epigenetic modification that usually silences gene expression (Newell-Price et al. 2000). Conversely, lower or no methylation has been associated with upregulated gene expression. Lower global methylation content across the human epigenome has been associated with hyperglycemia and the up-regulation of inflammatory genes in nondiabetics. However, some figures and Supplemental Material published in EHP articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehponline@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.
peripheral leukocytes from patients with type 2 diabetes mellitus (T2DM) (Luttmer et al. 2013; Nilsson et al. 2014). Yet, the role of methylation of specific genes related to inflammation in mediating the effects of PM on FBG has not been investigated.

In a repeated measure study of older men in the Greater Boston Area, we investigated the association between ambient PM$_{2.5}$ (PM with aerodynamic diameter ≤ 2.5 µm) concentrations at the participants’ addresses—estimated over different time windows up to 1 month before the visit—and FBG levels among nondiabetic participants. Using recently developed repeated measure mediation analysis, we further examined whether and to what extent PM$_{2.5}$ increased FBG through changes in blood leukocyte methylation in candidate inflammatory genes. We examined methylation of inflammatory cytokines (interferon gamma (IFN-γ) and interleukin-6 (IL-6)), intercellular adhesion molecule-1 (ICAM-1), and Toll-like receptor 2 (TLR-2). We hypothesized that higher PM$_{2.5}$ levels associated with increased FBG, and part of this association was mediated through methylation of inflammatory genes.

**Materials and Methods**

**Study Population**

The Normative Aging Study is a prospective longitudinal cohort established in 1963 by the U.S. Veterans Administration in the Greater Boston area (Power et al. 2011). Briefly, participants underwent examinations every 3–5 years. Self-administered questionnaires were collected at each visit providing information on sociodemographic characteristics, medical history, medications, and lifestyle. Blood samples were collected at each clinical visit, after an overnight fast and smoking abstinence (see Figure S2). Starting from year 2000, estimated concentrations of PM$_{2.5}$ were obtained from a hybrid spatiotemporal prediction model, as described in the next section. A total of 656 participants had complete information on PM$_{2.5}$ measurements from the prediction model, FBG, and blood leukocyte DNA methylation for at least one and up to four visits between 2000 and 2011 (median = 2, IQR = 1). We excluded 105 participants who, at their first visit, were clinically diabetic (FBG ≥ 126 mg/dL at the visit) and were taking diabetes medications. Therefore, our final study population included 551 participants. Fifty-two participants were diagnosed with diabetes during subsequent visits: For these individuals, we retained observations from the visit(s) before they were diagnosed with diabetes. One hundred and eighty-six participants came to just one clinical visit, 163 participants came to two clinical visits, and 202 participants came to three or more clinical visits. Participants provided signed informed consent at each visit. The study was approved by the Institutional Review Boards of the participating institutions.

**Air Pollution and Temperature**

We estimated PM$_{2.5}$ concentrations at each participant’s residential address using a hybrid land use regression and satellite-based model (Kloog et al. 2012; Madrigano et al. 2013). In brief, we utilized MODIS (Moderate Resolution Imaging Spectroradiometer) satellite-derived aerosol optical depth (AOD) measurements to predict daily PM$_{2.5}$ concentration levels at a 10 km spatial resolution. Daily AOD was calibrated using ground PM$_{2.5}$ measurements from 78 monitoring stations, land use regression and meteorological variables. To estimate PM$_{2.5}$ daily concentrations in each grid cell, we calibrated the AOD-PM$_{2.5}$ relationship using data from grid cells with both monitor and AOD values, using mixed models with random slopes for day and nested regions. In a later stage, we estimated exposures on days when AOD measures were not available (e.g., due to cloud cover or snow). Model performance was good with high out-of-sample 10-fold cross-validated $R^2$ (mean out-of-sample $R^2$ = 0.83 and 0.81 for days with and without available AOD data, respectively) (Kloog et al. 2011).

Temperature values were obtained through the national climatic data center. Only continuous operating stations that collected data on a daily basis were used (26 stations). Grid cells were matched to the closest weather station for meteorological variables.

**DNA Methylation Measurements**

Gene-specific DNA methylation was quantified on buffy-coat DNA using bisulfite polymerase-chain-reaction pyrosequencing (Yang et al. 2004). In the Normative Aging Study, we generated pyrosequencing-based methylation data for nine genes across pathways of inflammation (Bind et al. 2014). We chose to focus our analysis on the four inflammatory genes included in these pathways, including two inflammatory cytokines (IFN-γ and IL-6), intercellular adhesion molecule 1 (ICAM-1), and Toll-like receptor 2 (TLR-2). The choice of these four inflammatory genes was aligned with previous literature on PM and inflammation, as well as subclinical inflammation and the risk of T2DM (Bind et al. 2012; Blüher et al. 2002; Herder et al. 2013; Madrigano et al. 2010; van Eeden et al. 2001; Vossoughi et al. 2014). Specifically, IFN-γ and IL-6 encode for inflammatory cytokines that facilitate cell-to-cell communications in the inflammatory cascade, ICAM-1 encodes for a glycoprotein that is often expressed on the cell surface of endothelial cells and leukocytes, and TLR-2 encodes for a surface receptor protein, which recognizes conservative molecular patterns and serves as first-line defense in innate immunity. DNA methylation levels were measured for each of these genes at two to five CpG sites within the promoter region except for IL-6 methylation, which was measured within 500 base-pair downstream of the gene’s promoter region where nuclear respiratory factor-1 (NRF-1) binding sites are located (see Figure S1). The methylation of NRF-1 region is known to suppress IL-6 gene expression (Choi et al. 2004). We calculated and used the mean level of position-specific DNA methylation for each gene because they are highly correlated and are likely to share most of the same functional complexes and traits, and we assumed that mean methylation across the promoter region reflects regional epigenetic regulation (Bind et al. 2014).

**Fasting Blood Glucose Measurement**

Blood glucose levels were measured at each visit, after an overnight fast and were analyzed using the enzymatic hexokinase method. According to American Diabetes Association criteria, FBG less than 100 mg/dL corresponds to normal levels and the 100–125 mg/dL range to impaired fasting blood glucose (IFG), which in older individuals often progresses to diabetes over time. FBG larger than 125 mg/dL is defined as clinical diabetes (American Diabetes Association 2008, 2012).

**Statistical Analysis of the Main Association of PM$_{2.5}$ Levels with FBG**

We evaluated the association between PM$_{2.5}$ levels and FBG (modeled as a continuous dependent variable) using linear mixed-effects regression with subject-specific intercepts to account for the correlation among repeated FBG measurements within the same individual. Exposure variables included averages of PM$_{2.5}$ concentrations for 1-, 7-, and 28-day preceding each clinical visit; we considered each moving average in a separate regression model. Model estimates are expressed per interquartile range (IQR) increase in PM$_{2.5}$ concentration. In the models, we adjusted for the following covariates selected a priori: age (continuous), body mass index [BMI; weight (kg)/height (m)$^2$, continuous], race (white or others), regular patterns of physical activity (< 12 hr/week, ≥ 12 and < 30 hr/week, ≥ 30 hr/week), smoking status (never, former, or current smoker), cumulative pack-years of smoking (continuous), alcohol consumption (< 2 or ≥ 2 drinks/week), education level (high school diploma or less, college degree, or graduate degree), statin use (nonuser, statin user).
current user), temperature (continuous), and seasonality. Seasonality was modeled using Fourier series terms \( \cos(2\pi \times \text{doy}/365.25) \) and \( \sin(2\pi \times \text{doy}/365.25) \), where doy represents day of year. We checked the linearity assumptions of the continuous covariates using cubic splines and found no deviation from linear dose response.

In Equation 1, the main regression model took the general form:

\[
Y_{ij} = \beta_0 + u_i + \beta_1 X_{1ij} + \ldots + \beta_p X_{p pij} + \beta_{PM2,5,PM2.5} + \epsilon_{ij},
\]

where \( i \) corresponds to each participant, \( j \) to the visit; \( \beta_0 \) to the intercept for the population mean; \( u_i \) to the subject-specific random intercept; \( \beta_1 X_{1ij} \) to \( \beta_p X_{p pij} \) correspond to the covariates we selected \( \text{a priori} \). \( \beta_{PM2,5,PM2.5} \) corresponds to PM2.5 levels 1-, 7-, and 28-day prior to the clinical visits, respectively. \( Y_{ij} = 0 \) corresponds to PM2.5 levels 1-, 7-, or 28-day prior to the clinical visits, depending on the moving average used in each set of models. \( \epsilon_{ij} \) is the within-subject error term.

In a secondary analysis, we considered a dichotomized FBG5 variable for impaired fasting glucose (IFG) (categorized using the 0–100 mg/dL and 100–125 mg/dL ranges) as the outcome and evaluated the association between PM2.5 and the odds of IFG using a logistic regression model with generalized estimating equations (GEE) and empirical variance estimates to account for repeated measurements per subject.

In Equation 2, the logistic regression model took the general form:

\[
\logit \{ \Pr(Y_{ij} = 1) \} = \beta_0 + \beta_1 X_{1ij} + \ldots + \beta_p X_{p pij} + \beta_{PM2,5,PM2.5} + \epsilon_{ij},
\]

where \( i \) corresponds to each participant, \( j \) to the visit; \( \beta_0 \) to the intercept for the population mean. \( \beta_1 X_{1ij} \) to \( \beta_p X_{p pij} \) correspond to the covariates we selected \( \text{a priori} \). \( \beta_{PM2,5,PM2.5} \) corresponds to PM2.5 levels 1-, 7-, and 28-day prior to the clinical visits, respectively. \( Y_{ij} = 0 \) indicates subject \( i \) is not defined as IFG at visit \( j \); \( Y_{ij} = 1 \) indicates subject \( i \) is IFG at visit \( j \). To account for potential selection bias due to loss of follow-up, we repeated our analyses using inverse probability weighting. Specifically, in a logistic regression, we predicted the probability of coming to a subsequent visit based on covariates from the previous one, including age, BMI, regular patterns of physical activity, smoking status, pack-year smoked, alcohol consumption, education level, skin tone, temperature, seasonality, batch of methylation measurement, percentages of lymphocytes, and percentage of neutrophils. We assumed no unmeasured confounding for a) PM2.5-FBG relation, b) methylation-FBG relation, c) methylation-methylation relation, after fitting the linear mixed-effects models with subject-specific intercepts and controlling for \( C_1, C_2, \) and \( C_3 \). In addition, we also assumed that no methylation-FBG confounders would be affected by PM2.5 exposure.

Due to the longitudinal nature of the study, changes in FBG at one visit could potentially affect gene-specific methylation at the subsequent visit \( Y_{ij} \rightarrow M_{ij+1} \) (dotted arrow in Figure 1). FBG at one visit therefore may serve as a potential mediator-outcome confounder for the subsequent visit and may introduce bias in our estimates (Diggle et al. 2013). Therefore, in Equation 3, we tested the presence of an association between \( Y_{ij} \) and \( M_{ij+1} \), to check the assumption of time-varying confounding:

\[
M_{ij+1} = \alpha_0 + u_i + \alpha_1 Y_{ij} + \ldots + \alpha_p X_{p pij} + \alpha_{PM2,5,PM2.5} + \epsilon_{ij},
\]

where \( i \) corresponds to each participant and \( j \) to the visit; \( \alpha_0 \) to the intercept for the population mean; \( u_i \) to the subject-specific random intercept. \( \alpha_1 X_{1ij} \) corresponds to DNA methylation at subsequent visit. \( Y_{ij} \) corresponds to FBG measurement. \( \alpha_{PM2,5,PM2.5} \) corresponds to PM2.5 levels 1-, 7-, and 28-day prior to the clinical visits, respectively. \( \epsilon_{ij} \) is the within-subject error term.

**Mediation analysis.** We fitted two linear mixed-effects models with random intercepts simultaneously, one modeling the exposure-mediator association, and one modeling the mediator-outcome association (Bauer et al. 2006; Bind et al. 2010) (Equations 4 and 5):

\[
Y_{ij} = \gamma_0 + \gamma_0 X_{1ij} + \ldots + \gamma_p X_{p pij} + \gamma_{PM2,5,PM2.5} + \gamma u_i + \eta_{ij},
\]

where \( i \) corresponds to each participant and \( j \) to the visit; \( \gamma_0 \) and \( \gamma_0 \) to the intercept for the population mean; \( u_i \) to the subject-specific random intercept. \( \gamma_{PM2,5,PM2.5} \) corresponds to PM2.5 levels 1-, 7-, and 28-day prior to the clinical visits, in the exposure–mediator association, and \( \gamma u_i \) corresponds to DNA methylation in the mediator-outcome association. \( \eta_{ij} \) and \( \eta_{ij} \) are the within-subject error terms.

**Statistical Analysis of DNA Methylation and Mediation Analysis**

**Selection of mediators.** We hypothesized that associations of PM2.5 with FBG could be mediated through changes in gene-specific methylation of inflammatory biomarkers.

We considered the four inflammation genes—\( IFN-\gamma, IL-6, ICAM-1 \) and \( TLR-2 \)—separately in the mediation analysis. To approximate normality of the residuals, we used \( IFN-\gamma \) and \( IL-6 \) on their original scale and log-transformed \( ICAM-1 \) and \( TLR-2 \). For DNA methylation of a specific gene to be considered as a potential mediator, we tested the following criteria \( a \) if there was an association between exposure and mediator; and \( b \) if there was an association between mediator and outcome (Baron and Kenny 1986; Valeri and VanderWeele 2013). We also examined the presence of PM-mediator interactions and found no evidence of interactions that changed FBG levels.

**Underlying assumptions.** To obtain valid estimates of the natural indirect effects, we adjusted for potential exposure-outcome confounders (denoted as \( C_1 \)), exposure-mediator confounders (denoted as \( C_2 \)), and mediator-outcome confounders (denoted as \( C_3 \)), which included age, BMI, race, regular patterns of physical activity, smoking status, cumulative pack-years smoked, alcohol consumption, education level, skin tone, temperature, seasonality, patch of methylation measurement, percentages of lymphocytes, and percentage of neutrophils. We assumed no unmeasured confounding for a) PM2.5-FBG relation, b) methylation-FBG relation, c) methylation-methylation relation, after fitting the linear mixed-effects models with subject-specific intercepts and controlling for \( C_1, C_2, \) and \( C_3 \). In addition, we also assumed that no methylation-FBG confounders would be affected by PM2.5 exposure.

Due to the longitudinal nature of the study, changes in FBG at one visit could potentially affect gene-specific methylation at the subsequent visit \( Y_{ij} \rightarrow M_{ij+1} \) (dotted arrow in Figure 1). FBG at one visit therefore may serve as a potential mediator-outcome confounder for the subsequent visit and may introduce bias in our estimates (Diggle et al. 2013). Therefore, in Equation 3, we tested the presence of an association between \( Y_{ij} \) and \( M_{ij+1} \), to check the assumption of time-varying confounding:

\[
M_{ij+1} = \alpha_0 + u_i + \alpha_1 Y_{ij} + \ldots + \alpha_p X_{p pij} + \alpha_{PM2,5,PM2.5} + \epsilon_{ij},
\]

where \( i \) corresponds to each participant and \( j \) to the visit; \( \alpha_0 \) to the intercept for the population mean; \( u_i \) to the subject-specific random intercept. \( \alpha_1 X_{1ij} \) corresponds to DNA methylation at subsequent visit. \( Y_{ij} \) corresponds to FBG measurement. \( \alpha_{PM2,5,PM2.5} \) corresponds to PM2.5 levels 1-, 7-, and 28-day prior to the clinical visits, respectively. \( \epsilon_{ij} \) is the within-subject error term.

**Mediation analysis.** We fitted two linear mixed-effects models with random intercepts simultaneously, one modeling the exposure-mediator association, and one modeling the mediator-outcome association (Bauer et al. 2006; Bind et al. 2010) (Equations 4 and 5):

\[
Y_{ij} = \gamma_0 + \gamma_0 X_{1ij} + \ldots + \gamma_p X_{p pij} + \gamma_{PM2,5,PM2.5} + \gamma u_i + \eta_{ij},
\]

where \( i \) corresponds to each participant and \( j \) to the visit; \( \gamma_0 \) and \( \gamma_0 \) to the intercept for the population mean; \( u_i \) to the subject-specific random intercept. \( \gamma_{PM2,5,PM2.5} \) corresponds to PM2.5 levels 1-, 7-, and 28-day prior to the clinical visits, in the exposure–mediator association, and \( \gamma u_i \) corresponds to DNA methylation in the mediator-outcome association. \( \eta_{ij} \) and \( \eta_{ij} \) are the within-subject error terms.

**Figure 1.** Directed acyclic graph (DAG) for mediation analysis. PM2.5ij represents air pollution exposure for jth subject prior to j = jth visit; Mij represents gene-specific DNA methylation for ith subject at jth visit; Yij represents fasting blood glucose (FBG) concentrations for ith subject at jth visit. Cij represents exposure outcome confounders; Cij represents exposure mediator confounders; Ci represents mediator outcome confounders. Note: to be simplified, correlations between repeated measures of exposures (i.e., PM2.5ij and PM2.5ij-1), repeated measures of mediators (i.e., Mij and Mij-1) and repeated measures of confounders (i.e., Ci and Cij-1) are not shown in this DAG.
PM2.5 corresponds to the natural direct effect, and the natural indirect effect (also called “mediated” effect) is given by the product of $\beta_{PM2.5} \times \gamma_M$. The delta method was used to calculate the variance of the natural indirect effect, which corresponds to $\text{Var}(\gamma_M) \beta_{PM2.5}^2 + 2\text{Cov}(\beta_{PM2.5}, \gamma_M) \beta_{PM2.5} + \text{Var}((\beta_{PM2.5}) \gamma_M)$. Proportion mediated is calculated as the percentage of natural indirect effect over the sum of natural direct and natural indirect effect (i.e., $\frac{\beta_{PM2.5} \gamma_M}{\beta_{PM2.5} \gamma_M + \text{Var}(\beta_{PM2.5} \gamma_M)}$).

In sensitivity analysis, we tested the robustness of the study findings to the unmeasured confounding assumptions: a) we excluded participants who were current smokers to better control for residual confounding by smoking; b) we additionally controlled for total calorie intake and glyceric index to reduce potential mediator-outcome confounding from diet; c) we also restricted the analysis to participants with a C-reactive protein (CRP) level < 10 mg/L to partially remove potential effects from acute inflammation.

All analyses were conducted with SAS (version 9.3; SAS Institute Inc.), using PROC MIXED to fit the linear mixed effect models and PROC GENMOD to fit the GEE models.

Results

Descriptive Statistics

A total of 1,152 FBG measurements were collected from the 551 nondiabetic participants in this study. Table 1 describes participant characteristics at their first visit. Mean FBG concentration at the first visit was 99.8 ± 10.7 mg/dL (mean ± SD) and 18% of participants at the first visit had IFG (i.e., fasting blood glucose between 100 mg/dL and 126 mg/dL). Summary statistics of PM2.5 and temperature during the study period are presented in Table S1.

Main Association of PM2.5 Levels with FBG

PM2.5 levels were associated with increased FBG (Table 2). For an IQR increase in PM2.5 concentration in the previous 1-day (IQR = 5.73 μg/m³), 7-day (IQR = 4.25 μg/m³), and 28-day (IQR = 3.12 μg/m³) exposure windows were associated with OR equal to 1.13 (95% CI: 0.97, 1.33, p-value = 0.12), 1.27 (95% CI: 1.06, 1.52, p-value = 0.01) and 1.32 (95% CI: 1.10, 1.58, p-value = 0.003) for IFG, respectively. We also obtained similar estimates when we used inverse probability weighting to reduce potential selection bias (see Table S2).

DNA Methylation and Mediation Analysis

PM2.5 showed a negative association with ICAM-1 methylation, which, in turn, was negatively associated with FBG (Figure 2). Methylation of IFN-γ, IL-6 and TLR-2 showed no association with FBG. We also found no evidence of PM2.5-mediator interactions that changed FBG levels. Since a mediator needs to be associated with both the exposure and the outcome (Baron and Kenny 1986; Valeri and Vanderweele 2013), we conducted the analysis of mediation only for ICAM-1 methylation. We examined the correlations between ICAM-1 methylation and the other three genes and found no substantial correlations, which suggests that the separate analyses are fairly appropriate.

Table 1. Characteristics of the Normative Aging Study participants included in the analysis, 2000–2011.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Visit 1</th>
<th>Visit 2</th>
<th>Visit 3</th>
<th>Visit 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean ± SD</td>
<td>73.3 ± 6.9</td>
<td>75.6 ± 6.4</td>
<td>77.9 ± 5.9</td>
<td>78.5 ± 5.8</td>
</tr>
<tr>
<td>BMI (kg/m²), mean ± SD</td>
<td>27.8 ± 3.7</td>
<td>27.4 ± 3.7</td>
<td>27.1 ± 3.6</td>
<td>27.5 ± 3.9</td>
</tr>
<tr>
<td>Smoking status, n(%)</td>
<td>165 (30%)</td>
<td>112 (21%)</td>
<td>71 (24%)</td>
<td>14 (33%)</td>
</tr>
<tr>
<td>Metabolic equivalent of task, n(%)</td>
<td>353 (64%)</td>
<td>222 (61%)</td>
<td>129 (61%)</td>
<td>26 (60%)</td>
</tr>
<tr>
<td>Low (≤ 12 hr/week)</td>
<td>111 (20%)</td>
<td>89 (25%)</td>
<td>47 (22%)</td>
<td>10 (22%)</td>
</tr>
<tr>
<td>Medium (12–30 hr/week)</td>
<td>87 (16%)</td>
<td>52 (14%)</td>
<td>35 (17%)</td>
<td>7 (16%)</td>
</tr>
<tr>
<td>Two or more drinks per day, n(%)</td>
<td>102 (19%)</td>
<td>69 (19%)</td>
<td>35 (17%)</td>
<td>4 (9%)</td>
</tr>
<tr>
<td>Education, n(%)</td>
<td>13–16 years: 253 (46%)</td>
<td>163 (45%)</td>
<td>98 (46%)</td>
<td>24 (56%)</td>
</tr>
<tr>
<td>Statin use, n(%)</td>
<td>196 (36%)</td>
<td>180 (50%)</td>
<td>122 (58%)</td>
<td>24 (56%)</td>
</tr>
<tr>
<td>Fasting blood glucose (mg/dL)</td>
<td>99.9 ± 10.7</td>
<td>99.9 ± 10.0</td>
<td>99.0 ± 10.5</td>
<td>99.0 ± 11.5</td>
</tr>
<tr>
<td>Blood IFN-γ methylation, mean ± SD</td>
<td>84.3 ± 5.9</td>
<td>85.0 ± 4.8</td>
<td>85.3 ± 4.7</td>
<td>85.3 ± 5.5</td>
</tr>
<tr>
<td>Blood IL-6 methylation, mean ± SD</td>
<td>43.6 ± 10.3</td>
<td>43.5 ± 10.2</td>
<td>43.8 ± 10.3</td>
<td>42.8 ± 11.5</td>
</tr>
<tr>
<td>Blood ICAM methylation, mean ± SD</td>
<td>4.3 ± 1.8</td>
<td>3.9 ± 1.2</td>
<td>4.4 ± 1.2</td>
<td>4.8 ± 1.9</td>
</tr>
<tr>
<td>Blood TLR2 methylation, mean ± SD</td>
<td>3.1 ± 1.3</td>
<td>3.0 ± 1.4</td>
<td>2.5 ± 1.4</td>
<td>1.9 ± 1.0</td>
</tr>
</tbody>
</table>

Note: Cohort participants with diabetes were excluded.

Table 2. Estimated change (95% CI) in fasting blood glucose level (mg/dL) per interquartile range increase in PM2.5 concentration averaged over the corresponding time window before each visit.

<table>
<thead>
<tr>
<th>PM2.5 concentration</th>
<th>Participants</th>
<th>Observations</th>
<th>PM2.5 interquartile range (IQR)</th>
<th>Estimated change (95% CI) in FBG per IQR increase in PM2.5 concentrations</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-day moving average</td>
<td>551</td>
<td>1,152</td>
<td>5.71 μg/m³</td>
<td>0.57 (0.02, 1.11)</td>
<td>0.04</td>
</tr>
<tr>
<td>7-day moving average</td>
<td>551</td>
<td>1,152</td>
<td>4.28 μg/m³</td>
<td>1.02 (0.41, 1.63)</td>
<td>0.001</td>
</tr>
<tr>
<td>28-day moving average</td>
<td>551</td>
<td>1,152</td>
<td>3.09 μg/m³</td>
<td>0.89 (0.32, 1.47)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, and seasonality. Participants with diabetes were excluded.

Table 3. Odds ratio (95% CI) of impaired fasting blood glucose (IFG) per interquartile range increase in PM2.5 concentration averaged over the corresponding time window before each visit.

<table>
<thead>
<tr>
<th>PM2.5 concentration</th>
<th>Participants</th>
<th>Observations</th>
<th>PM2.5 interquartile range (IQR)</th>
<th>Odds ratio (95% CI) of IFG per IQR increase in PM2.5 concentrations</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-day moving average</td>
<td>551</td>
<td>1,152</td>
<td>5.73 μg/m³</td>
<td>1.13 (0.97, 1.33)</td>
<td>0.12</td>
</tr>
<tr>
<td>7-day moving average</td>
<td>551</td>
<td>1,152</td>
<td>4.25 μg/m³</td>
<td>1.27 (1.06, 1.52)</td>
<td>0.01</td>
</tr>
<tr>
<td>28-day moving average</td>
<td>551</td>
<td>1,152</td>
<td>3.12 μg/m³</td>
<td>1.32 (1.10, 1.58)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Note: IFG is defined as a fasting blood glucose level > 100mg/dL and < 126 mg/dL. Results from GEE models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, and seasonality. Participants with diabetes were excluded.
levels among nondiabetic individuals, as well as with higher odds of IFG. We also observed significant associations of lower blood ICAM-1 methylation, which is expected to up-regulate the expression of ICAM-1 in blood leukocytes, with both higher PM$_{2.5}$ levels and higher FBG levels.

Our study is consistent with previous epidemiology studies indicating that ambient PM is associated with metabolic dysregulation (Andersen et al. 2012; Brook et al. 2008; Eze et al. 2015; Fuett et al. 2011; Rajagopalan and Brook 2012). Nevertheless, most previous studies either focused on T2DM or evaluated blood glucose over its entire range, including individuals with diabetes. The nondiabetic and pre-diabetic population represents an ideal target for primary prevention, which, however, has been understudied in air pollution research. Our findings of PM$_{2.5}$ associations with FBG in this group may help identify individuals who are particularly susceptible to changes in FBG in the nondiabetic range. It is important to notice that the small changes in FBG resulted in significant odds of IFG, owing to the fact that many people are very close to 100 mg/dL in this older population, and a very small change in FBG may be enough to pass the threshold.

Our mediation analysis suggests that ICAM-1 methylation in blood leukocytes served as a mediator of the association between PM$_{2.5}$ and FBG, and we observed significant mediated effect at 28-day exposure time window. Our finding of higher concentrations of PM$_{2.5}$ associated with lower methylation of the ICAM-1 gene, which is expected to result in higher ICAM-1 expression, is consistent with previous literature indicating that elevated concentrations of PM are associated with an increase in expression of endothelial markers (Bind et al. 2012; Madrignano et al. 2010; O’Neill et al. 2007; Rückerl et al. 2006). The ICAM-1 glycoprotein is responsible for leukocyte adhesion, homing, and transmigration during inflammatory responses (Rahman and Fazal 2009). Exposure to PM$_{2.5}$ may cause local inflammation in the lungs and promote circulating leukocytes in blood to transmigrate to the target tissue through the up-regulation of adhesion molecules on the endothelial cell surface. Recent observational and intervention studies have linked elevated concentrations of plasma endothelial adhesion molecules with markers of insulin resistance and increased risk of T2DM (Blüher et al. 2002; Hak et al. 2001; Meigs et al. 2004). The ICAM-1 glycoprotein may facilitate migration of leukocytes from the blood to the adipose tissue (Mendez et al. 2013; Rao et al. 2015; Sun et al. 2009), which could result in local inflammation and subsequently insulin resistance. Alternatively, the ICAM-1 glycoprotein may also facilitate leukocyte transmigration to the pancreas.

Discussion
In the present study, we showed that PM$_{2.5}$ concentrations estimated at the participants’ address were associated with higher FBG

Table 4. Mediation effect investigating whether blood ICAM-1 methylation mediates the association between air pollution and fasting blood glucose level. Indirect effect represents the mediated effect through the ICAM-1 methylation pathway.

<table>
<thead>
<tr>
<th>PM$_{2.5}$ concentration</th>
<th>Exposure to mediator association ($\beta_{MB_{i,j}}$)</th>
<th>Mediator to outcome association ($\gamma_{i,j}$)</th>
<th>Mediated effect of ICAM-1 methylation</th>
<th>Proportion mediated</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-day moving average</td>
<td>0.004 (0.008, 0.008)</td>
<td>-2.65 (-4.41, -0.89)</td>
<td>-0.01 (-0.02, 0.004)</td>
<td>1%</td>
</tr>
<tr>
<td>7-day moving average</td>
<td>-0.004 (-0.007, 0.006)</td>
<td>-2.69 (-4.45, -0.93)</td>
<td>0.001 (-0.02, 0.02)</td>
<td>1%</td>
</tr>
<tr>
<td>28-day moving average</td>
<td>-0.01 (-0.02, -0.004)</td>
<td>-2.47 (-4.23, -0.72)</td>
<td>0.03 (0.0001, 0.06)</td>
<td>9%</td>
</tr>
</tbody>
</table>

Note: Estimates correspond to 1-μg/m$^3$ increase in PM$_{2.5}$ concentration.

Results from linear mixed-effects regression models accounting for correlation across multiple visits and adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, temperature, seasonality, batch effect, % of lymphocytes, and % of neutrophils. Participants with diabetes were excluded.

*Proportion mediated cannot be estimated in this case because $\beta_{MB_{i,j}}$ and $\gamma_{i,j}$ have opposite signs.

Figure 2. Inflammatory candidate gene methylation mediator model of the relationship between PM$_{2.5}$ concentration and fasting blood glucose level.

ICAM-1 mean and TLR-2 mean DNA methylation is log-normally distributed and IL-6 and IFN-γ are normally distributed. $\beta$ is the coefficient of the independent variable (PM$_{2.5}$ 28-day moving average) when regressing the mediator (candidate gene methylation) on the independent variable, $\gamma$ is the coefficient of the mediator when regressing the dependent variable (FBG) on both the independent variable and the mediator. Results from regression models are adjusted for age, BMI, race, regular patterns of physical activity, smoking status, pack-years smoked, alcohol consumption, education level, statin use, batch effects, percentage of lymphocytes, and percentage of neutrophils. Participants with diabetes were included.
(Yagi et al. 1995), which could affect beta-cell function and result in impaired insulin secretion. Although the observed association was relatively modest, our estimates were comparable to many other studies evaluating the association between ambient air pollution and DNA methylation with similar exposure levels (Guo et al. 2014; Madrigano et al. 2011). DNA methylation is measured as a percentage, which indicates the proportion of cells, or more accurately of haploid genomes, which show methylation at the sequence being analyzed. The differences in DNA methylation reported in our study are related to the presence upon PM$_{2.5}$ exposure of higher numbers of circulating blood cells within no methylation at the ICAM-1 promoter. Further research is needed to determine whether these cells correspond to a specific leukocyte population with known function and their potential roles in relation to PM$_{2.5}$ effects.

Conversely, methylation on the cytokine genes investigated (i.e., IFN-$\gamma$ and IL-6) was not implicated as a mediator in this study. Many pro- and anti-inflammatory cytokines act in concert to trigger the inflammatory cascade. Future research may expand the number of inflammatory cytokines investigated and examine their joint effects.

One interesting aspect of the human methylome is that it exhibits both dynamic and static patterns. For instance, methylation of imprinted genes and genes drives tissue lineage commitment and differentiation that is established during embryogenesis and persists through life (Li et al. 1993): on the other hand, DNA methylation of inflammatory genes may change rapidly after environmental insults (Guo et al. 2014; Hou et al. 2014; Tarantini et al. 2009). Our results are consistent with the dynamic nature of methylation levels in inflammatory pathways, which allows for fine tuning of inflammatory responses. Nevertheless, we are aware that differences in DNA methylation do not necessarily translate into gene expression changes. DNA methylation is only one of the regulatory machineries that control gene expression. Other regulatory mechanisms, such as transcription factor activation, histone modification, chromatin remodeling and RNA silencing may also contribute to regulation of gene expression at various stages.

This study has a number of strengths. We estimated concentrations of PM$_{2.5}$ at the residential address for each participant using a state-of-art hybrid model. Estimates from this hybrid model serve as better surrogates relative to the standard use of data from monitoring stations for each participant’s actual exposure, and limit exposure misclassifications. We conducted analyses both on FBG as a continuous variable and on IFG, a dichotimized variable constructed using a well-established cutoff for preclinical alterations in glucose metabolism. These two sets of analyses produced highly consistent results. We used repeated measures, which were accounted for using linear mixed-effects models with subject specific intercepts for FBG and generalized linear equations (GEE) model with empirical variance for IFG. We conducted mediation analysis as a novel approach for DNA methylation studies. This approach is particularly useful for investigating effects due to environmental exposures.

While Mendelian randomization—a method that relies on genotype data used as instrumental variables—is often proposed to identifying epigenetic mediation, this approach cannot be used for external risk factors, such as PM, which, due to their nature, are not associated with the participants’ genetic sequences (such as single nucleotide polymorphisms) (Relton and Davey Smith 2012).

Finally, we measured DNA methylation in candidate genes by pyrosequencing, which yields high precision (Liao and Guo 2007).

Our study has a few notable limitations. We focused on correlations among DNA methylation and FBG at the same clinical visit. It is therefore difficult to disentangle the temporal relationship between DNA methylation and FBG concentrations. However, when we examined the effect of DNA methylation on FBG for the subsequent visit (i.e., $ICAM-I_{ij} \geq FBG_{ij+1}$), we did not find any association. We also found no association of FBG on ICAM-1 methylation at the subsequent visit (i.e., $FBG_{3} \geq \text{ICAM-I}_{3+1}$). These analyses confirm that the effects we observed represent short- and medium-term responses to PM$_{2.5}$ and are not persistent over the 3–5 years between the medical visits. To obtain valid estimates for the natural indirect effects, we made the following assumptions: a) no-unmeasured confounding between PM$_{2.5}$ concentration and FBG levels, b) no-unmeasured confounding between PM$_{2.5}$ concentration and methylation, c) no-unmeasured confounding between methylation and FBG levels, and d) no methylation-FBG confounders affected by the exposure. We conducted a number of sensitivity analyses to test the robustness to the no-unmeasured confounding assumptions. We excluded participants who were current smokers to better control for residual confounding by smoking: we additionally controlled for total calorie intake and glycemic index, to limit potential confounding from diet; we also restricted the analysis to participants with a CRP level $< 10$ mg/L, to partially remove potential effect from acute inflammation. In addition, we assessed the validity of assumption $d$ either based on subject knowledge or empirically (Bind et al. 2016). From previous subject knowledge, we assumed that PM$_{2.5}$ concentration would not affect participants’ age, race, smoking status, and statin use, as well as the batch of methylation measurements and seasonality. We also tested whether PM$_{2.5}$ concentration would influence participant’s BMI, regular patterns of physical activity, percentage of lymphocytes, and percentage of neutrophils, by regressing PM$_{2.5}$ concentration on each of these potential confounders. None of the above mediator–outcome confounders in current analysis were affected by the exposure (p-values were $0.10, 0.54, 0.50$, and $0.78$, respectively). Another limitation of our study is the potential for measurement error both in the exposure and FBG. However, we expect both measurement errors to be nondifferential and therefore to attenuate—rather than to cause—the observed significant associations.

**Conclusion**

In conclusion, we found that PM$_{2.5}$ concentrations are associated with higher FBG level, and this association was in part mediated through ICAM-1 gene methylation, particularly at the longer (28-day) moving average investigated. Our study demonstrates a novel approach of mediation analysis in epigenetic studies and highlights a mediating role of ICAM-1 gene methylation in air pollution–associated abnormal glucose metabolism. While the proportion mediated by ICAM-1 methylation alone is relatively modest, methylation of other genes not investigated in this study, independently or in combination with ICAM-1 methylation, may mediate larger proportions of PM$_{2.5}$ effects. Future epigenome-wide studies are needed to determine the extent to which DNA methylation contributes to mediate environmental effects on human metabolism.

**References**


