Leukocyte Traits and Exposure to Ambient Particulate Matter Air Pollution in the Women’s Health Initiative and Atherosclerosis Risk in Communities Study

Rahul Gondalia,1 Katelyn M. Holliday,1,2 Antoine Baldassari,1 Anne E. Justice,1,3 James D. Stewart,1 Duanping Liao,4 Jeff D. Yanosky,4 Stephanie M. Engel,1 Kristina M. Jordahl,5 Parveen Bhatti,6 Steve Horvath,6,7 Themistocles L. Assimes,8 James S. Pankow,7 Ellen W. Demerath,7 Weihua Guan,9 Myriam Fornage,11 Jan Bressler,12 Kari E. North,1,13 Karen N. Conneely,14 Yun Li,15,16,17 Lifang Hou,18,19 Andrea A. Baccarelli,20 and Eric A. Whitsel1,21

Department of Medicine, University of North Carolina Gillings School of Global Public Health, Chapel Hill, North Carolina
2Department of Community and Family Medicine, Duke University School of Medicine, Durham, North Carolina
3Geisinger Health System, Danville, Pennsylvania
4Division of Epidemiology, Department of Public Health Sciences, Pennsylvania State University College of Medicine, Hershey, Pennsylvania
5Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington
6Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California
7Department of Biostatistics, School of Public Health, University of California, Los Angeles, Los Angeles, California
8Department of Medicine, Stanford University School of Medicine, Stanford, California
9Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota
10Division of Biostatistics, University of Minnesota, Minneapolis, Minnesota
11Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, Texas
12Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, Texas
13Carolina Center for Genome Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina
14Department of Human Genetics, Emory University School of Medicine, Atlanta, Georgia
15Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina
16Department of Biostatistics, University of North Carolina Gillings School of Global Public Health, Chapel Hill, North Carolina
17Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina
18Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois
19Center for Population Epigenetics, Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, Illinois
20Laboratory of Environmental Epigenetics, Departments of Environmental Health Sciences and Epidemiology, Columbia University Mailman School of Public Health, New York, New York
21Department of Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

BACKGROUND: Inflammatory effects of ambient particulate matter (PM) air pollution exposures may underlie PM-related increases in cardiovascular disease risk and mortality, although evidence of PM-associated leukocytosis is inconsistent and largely based on small, cross-sectional, and/or unrepresentative study populations.

OBJECTIVES: Our objective was to estimate PM–leukocyte associations among U.S. women and men in the Women’s Health Initiative and Atherosclerosis Risk in Communities study (n = 165,675).

METHODS: We based the PM–leukocyte estimations on up to four visits per participant, at which peripheral blood leukocytes and geocoded address-specific concentrations of PM ≤ 10, ≤ 2.5, and 2.5–10 μm in diameter (PM2.5, and PM10, respectively) were available. We multiply imputed missing data using chained equations and estimated PM–leukocyte count associations over daily to yearly PM exposure averaging periods using center-specific, linear, mixed, longitudinal models weighted for attrition and adjusted for sociodemographic, behavioral, meteorological, and geographic covariates. In a subset of participants with available data (n = 8,457), we also estimated PM–leukocyte proportion associations in compositional data analyses.

RESULTS: We found a 12 cells/μL (95% confidence interval: −9, 33) higher leukocyte count, a 1.2% (0.6%, 1.8%) higher granulocyte proportion, and a −1.1% (−1.9%, −0.3%) lower CD8+ T-cell proportion per 10-μg/m3 increase in 1-month mean PM2.5. However, shorter-duration PM10 exposures were inversely and only modestly associated with leukocyte count.

DISCUSSION: The PM2.5–leukocyte estimates, albeit imprecise, suggest that among racially, ethnically, and environmentally diverse U.S. populations, sustained, ambient exposure to fine PM may induce subclinical, but epidemiologically important, inflammatory effects. https://doi.org/10.1289/EHP5360

Introduction

Exposures to airborne particulate matter (PM) ≤ 10, ≤ 2.5 and between 2.5 and 10 μm in diameter (PM10, PM2.5, and PM2.5–10, respectively) can trigger inflammatory responses that involve the release and hematogenous redistribution of leukocytes (Pope et al. 2016; Tan et al. 2000; Terashima et al. 1997a). Such responses may be key to the pathophysiology underpinning established associations between ambient PM, cardiovascular (CVD) disease risk, and mortality (Brook et al. 2010; Chi et al. 2016a; Di et al. 2017; Miller et al. 2007; Parker et al. 2018). However, evidence of PM-associated leukocytosis is inconsistent and mostly based on small studies and panels with limited generalizability (Brook et al. 2009; Dubowsky et al. 2006; Emmerechts et al. 2012; Gho et al. 2003; Gong et al. 2004; Huang et al. 2014; Jacobs et al. 2010; Mills et al. 2005, 2007; Pope et al. 2004, 2016; Riediker 2007; Salvi et al. 1999; Steenhouf et al. 2014; Törnqvist et al. 2007).

In larger, community- and population-based studies, short-duration PM10–leukocyte count associations are similarly inconsistent (Liao et al. 2005; Schwartz 2001; Seaton et al. 1999; Steinviel et al. 2008), although longer-duration PM10– and PM2.5–leukocyte count associations tend to be positive in published...
cross-sectional and longitudinal studies (Chen and Schwartz 2008; Chuang et al. 2011; Viehmann et al. 2015). Moreover, associations between short- and longer-term PM exposures and leukocyte count and its differential composition have not been thoroughly evaluated while controlling for known relationships among leukocyte traits (count and component proportions).

Associations between ambient PM exposures and leukocyte traits could nevertheless lend support to the hypothesized role of inflammation in PM-related pathogenesis. Furthermore, their magnitude would provide insight into PM associations with leukocyte-derived biomarkers such as DNA methylation (DNAm), a heritable but dynamic epigenetic modification that can influence gene expression. Indeed, epidemiologic studies often rely on peripheral blood leukocytes as a source of DNA for DNAm assays given the relative ease with which they are collected and archived in large populations (McCullough et al. 2017; Zhong et al. 2016). Because DNAm and other epigenetic biomarkers (Beaulieu et al. 2017) differ among leukocyte subtypes [e.g., granulocytes vs. monocytes (Houseman et al. 2012; Jaffe and Irizarry 2014)], leukocyte composition may plausibly mediate their associations with environmental exposures.

To expand on prior work evaluating PM–leukocyte count associations, and to address the limitations of studies examining PM–leukocyte compositional associations, we estimated associations of leukocyte traits with short- to longer-duration exposures to ambient PM$_{2.5}$, PM$_{10}$, and PM$_{2.5-10}$ in large, multiracial/ethnic, and geographically diverse United States populations enrolled in the Women’s Health Initiative (WHI) and the Atherosclerosis Risk in Communities (ARIC) study.

**Methods**

**Study Populations**

The WHI is a multicenter prospective study of risk factors for CVD, breast/colorectal cancer, and osteoporotic fractures (Women’s Health Initiative Study Group 1998, Anderson et al. 2003). From forty clinical centers throughout the United States, postmenopausal women aged 50–79 years of age were either randomized in the Clinical Trials (CT; $n = 68,132$) or enrolled in the Observational Study (OS; $n = 93,676$) between 1993 and 1998. The WHI CT included three interventions: (a) hormone therapy (i.e., estrogen with or without progesterin vs. placebo), (b) calcium and vitamin D supplementation (vs. placebo), and (c) dietary modification (vs. usual diet). The WHI OS (Women’s Health Initiative Study Group 1998, Anderson et al. 2003) recruited participants interested in the dietary modification or hormone therapy trials of the WHI CT but were otherwise ineligible, unwilling, or unresponsive to a direct invitation.

The WHI CT and OS participants completed a baseline screening visit, at which fasting blood and other demographic, socioeconomic, behavioral, and medical information was collected by trained and certified staff. The present study additionally included WHI CT participant data from triennial follow-up visits 3 and 6 y after randomization (Annual Visits 3 and 6) and WHI OS participant data 3 y after enrollment (Annual Visit 3), at which fasting blood was redrawn.

The ARIC study is a prospective epidemiologic study of atherosclerosis and CVD in four U.S. communities: Washington County, Maryland; Forsyth County, North Carolina; selected suburbs of Minneapolis, Minnesota; and Jackson, Mississippi (ARIC Investigators 1989). Participants were selected as a community-stratified probability sample of 15,792 mostly African- and European-American men and women 45–64 years of age and participated in a baseline exam (Visit 1; 1987–1989) at which fasting blood and other demographic, socioeconomic, behavioral, and medical information was collected by trained and certified staff.

The present study also included participant data from up to three triennial follow-up visits 3, 6, and 9 y after enrollment (Visits 2–4, 1990–1998) during which fasting blood was redrawn.

Leukocyte composition analyses were conducted in five WHI and ARIC subpopulations with available DNAm data (see Table S1). The three WHI subpopulations included $a$) Ancillary Study 315 (WHI-EMPC; $n = 2,200$) (Whitsel 2018), $b$) Broad Agency Announcement 23 (WHI-BAAA23; $n = 1,988$) (Assimes et al. 2018), and $c$) Ancillary Study 311 (WHI-AS311; $n = 860$) (Bhatti 2018). WHI-EMPC, also known as Epigenetic Mechanisms of PM-Mediated CVD Risk, is a study of epigenetic mechanisms underlying associations between PM and CVD within randomly selected WHI CT participants at the screening visit, Annual Visit 3, or Annual Visit 6. WHI-BAAA23, also known as Integrative Genomics and Risk of CHD and Related Phenotypes in the Women’s Health Initiative, is a case–control study of coronary heart disease. By design, WHI-BAAA23 oversampled African Americans and Hispanic/Latino Americans and required all participants to have undergone genome-wide genotyping and profiling of seven CVD biomarkers. DNAm was measured in blood collected at the screening visit, before the incidence of coronary heart disease. WHI-AS311, also known as the Bladder Cancer and Leukocyte Methylation study, is a nested case–control study of bladder cancer. Bladder cancer cases were matched to controls based on enrollment year, age at enrollment, follow-up time, and DNAm extraction method. DNAm was measured in blood collected at the screening visit, before the incidence of bladder cancer. The two ARIC subpopulations included 2,796 African Americans from Forsyth County or Jackson (ARIC-AA) with DNA and 1,139 European Americans from Forsyth County, Minneapolis, or Washington County (ARIC-EA) with cerebral magnetic resonance imaging data (Mosley et al. 2005) all at Visits 2 (1990–1992) or 3 (1993–1995) (see Figure S1).

**Leukocyte Counts and Composition**

Leukocyte counts were measured among WHI CT participants at the screening visit, among OS participants at the screening visit and Annual Visit 3, and among ARIC participants at Visits 1–2 on automated cell counters at local laboratories following standard quality assurance procedures (Papp et al. 1989). Leukocyte counts were remeasured among ARIC participants in Washington County at Visits 3–4 and in Forsyth County at Visit 4. Table 1 displays the number of included participants with leukocyte count data, by study and visit. Established associations between leukocyte count, demographic, and clinical variables in WHI and ARIC have been reported by others (Margolis et al. 2005; Nieto et al. 1992).

Leukocyte composition [i.e., the proportions of CD8$^+$ T cells, CD4$^+$ T cells, natural killer (NK) cells, B cells, monocytes, and granulocytes] were validly estimated (Houseman et al. 2012) among a subset of WHI and ARIC participants with DNAm data using methods that leverage differentially methylated regions [i.e., stably methylated CpG sites within, but variably methylated CpG sites among leukocyte cell types (Houseman et al. 2012; Koestler et al. 2013)]. Table S2 displays the number of included participants with leukocyte composition data, by subpopulation.

**Particulate Matter Exposure Estimation**

The study focused on PM$_{2.5}$, PM$_{10}$, and (coarse) PM$_{2.5-10}$, the first two of which are regulated under the Clean Air Act by the U.S. Environmental Protection Agency (EPA) (U.S. EPA 2017). PM exposures were based on either daily or monthly estimation methods. Daily mean concentrations (in micrograms per cubic meter) of PM$_{10}$ were spatially estimated at all geocoded participant addresses.
### Table 1. Characteristics of \( n = 165,675 \) participants with leukocyte count data before imputation, Women’s Health Initiative (1993–2002) and Atherosclerosis Risk in Communities (1986–1998) study.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>WHI screening visit and ARIC visit 1</th>
<th>WHI</th>
<th>ARIC</th>
<th>WHI and ARIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( n = 159,162 )</td>
<td>Screening visit ( n = 144,744 ) Annual visit 3</td>
<td>Visit 1 ( n = 14,418 ) Visit 2 ( n = 13,000 ) Visit 3</td>
<td>Visit 4 ( n = 5,433 ) Percentage imputed of 285,548 observations (%)</td>
</tr>
<tr>
<td>Male [( n (%) )]</td>
<td>6,563 (4.1)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>6,563 (45.5)</td>
</tr>
<tr>
<td>Age [y (mean ± SD)]</td>
<td>62.3 ± 7.6</td>
<td>63.2 ± 7.2</td>
<td>66.5 ± 7.3</td>
<td>54.2 ± 5.8</td>
</tr>
<tr>
<td>Race/ethnicity [( n (%) )]</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaskan Native</td>
<td>658 (0.4)</td>
<td>647 (0.4)</td>
<td>315 (0.4)</td>
<td>11 (0.1)</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>1,611 (1.1)</td>
<td>1,018 (1.3)</td>
<td>32 (0.2)</td>
<td>29 (0.2)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>15,809 (10.0)</td>
<td>11,990 (8.3)</td>
<td>5,675 (7.4)</td>
<td>3,819 (26.5)</td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>5,967 (3.8)</td>
<td>5,967 (4.1)</td>
<td>2,681 (3.5)</td>
<td>2 (0.1)</td>
</tr>
<tr>
<td>Other</td>
<td>1,353 (0.9)</td>
<td>1,353 (0.9)</td>
<td>740 (1.0)</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>White (not of Hispanic origin) or European American</td>
<td>123,400 (84.0)</td>
<td>122,844 (85.1)</td>
<td>66,457 (86.4)</td>
<td>10,556 (73.2)</td>
</tr>
<tr>
<td>Education [( n (%) )]</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school education or lower</td>
<td>40,473 (25.6)</td>
<td>32,358 (22.5)</td>
<td>15,677 (20.5)</td>
<td>8,115 (56.4)</td>
</tr>
<tr>
<td>More than high school</td>
<td>117,654 (74.4)</td>
<td>111,377 (77.5)</td>
<td>60,818 (79.5)</td>
<td>6,277 (43.6)</td>
</tr>
<tr>
<td>Smoking status [( n (%) )]</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>78,794 (50.1)</td>
<td>72,760 (50.9)</td>
<td>37,749 (51.1)</td>
<td>6,032 (41.9)</td>
</tr>
<tr>
<td>Former</td>
<td>64,941 (41.2)</td>
<td>60,314 (42.2)</td>
<td>32,708 (44.2)</td>
<td>4,627 (32.1)</td>
</tr>
<tr>
<td>Current</td>
<td>13,564 (8.6)</td>
<td>9,821 (6.9)</td>
<td>3,465 (4.7)</td>
<td>3,743 (26.0)</td>
</tr>
<tr>
<td>Alcohol use [( n (%) )]</td>
<td>1.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>18,683 (11.8)</td>
<td>15,101 (10.5)</td>
<td>6,807 (9.1)</td>
<td>3,582 (24.9)</td>
</tr>
<tr>
<td>Former</td>
<td>28,972 (18.3)</td>
<td>26,274 (18.3)</td>
<td>15,040 (20.1)</td>
<td>2,698 (18.8)</td>
</tr>
<tr>
<td>Current</td>
<td>110,366 (69.8)</td>
<td>102,289 (71.2)</td>
<td>52,866 (70.8)</td>
<td>8,077 (56.3)</td>
</tr>
<tr>
<td>Body mass index [kg/m² (mean ± SD)]</td>
<td>28.0 ± 5.8</td>
<td>28.9 ± 5.9</td>
<td>27.4 ± 5.7</td>
<td>27.7 ± 5.3</td>
</tr>
<tr>
<td>Physical activity [MET-h/week (mean ± SD)]</td>
<td>12.3 ± 13.7</td>
<td>12.5 ± 13.7</td>
<td>13.7 ± 14.6</td>
<td>10.2 ± 12.7</td>
</tr>
<tr>
<td>Neighborhood SES ( z )-score sum</td>
<td>-0.1 ± 5.4</td>
<td>-0.1 ± 5.4</td>
<td>0.2 (5.3)</td>
<td>0.0 (5.4)</td>
</tr>
<tr>
<td>Leukocyte count [cells/lL]</td>
<td>5,908 ± 1,553</td>
<td>5,882 ± 1,529</td>
<td>5,794 ± 1,500</td>
<td>6,076 ± 1,761</td>
</tr>
</tbody>
</table>

Note: ARIC, Atherosclerosis Risk in Communities; SD, standard deviation; SES, socioeconomic status; WHI, Women’s Health Initiative.

aWHI Observational Study participants only.
bParticipants from Washington County only.
cParticipants from Forsyth County (46%) or Washington County (54%).
dARIC recruitment and data collection occurred before the National Institutes of Health required collection of information about Hispanic/Latino ethnicity.
(Whit sel et al. 2004, 2006) using U.S. EPA Air Quality System (AQS) data and national-scale, log-normal ordinary kriging (Liao et al. 2006, 2007). For each participant, daily mean concentrations of PM10 were averaged over 2 and 7 d prior to and including the day of the study visit.

Geocoded participant address-specific monthly mean concentrations (in micrograms per cubic meter) of PM10 and PM2.5 were spatiotemporally estimated using generalized additive mixed models and geographic information system–based predictors. Because U.S. EPA AQS monitoring data for PM2.5 were not widely available until 1999, spatiotemporal estimation also involved the log-transformed ratio of PM2.5 to predicted PM10 between 1987 and 1999 (Yanosky et al. 2014). Monthly mean concentrations were averaged over the 12 months prior to and including examination months to obtain annual means. PM2.5–10 concentrations for 1- and 12-month means were defined as the monthly differences between PM10 and PM2.5 concentrations.

Covariates
Demographic, socioeconomic, behavioral, and medical covariates included study center, visit, self-identified race/ethnicity, age (in years), individual-level education (high school education or lower, more than high school), neighborhood socioeconomic status (Diez Roux et al. 2001), smoking status (current, former, never), alcohol use (current, former, never), measured body mass index (BMI; in kilograms per squared meters), total energy expenditure [metabolic equivalent of task (MET)-hours/week], mean temperature (in degrees Celsius), mean dew point (in degrees Celsius), mean barometric pressure (in kilopascals), season (using sine/cosine functions) (Stolwijk et al. 1999), and to control for longer-term temporal trends, an interval-scale variable for calendar date. Race/ethnicity and individual-level education were self-reported at baseline. Smoking status, alcohol use, BMI, and total energy expenditure were evaluated at each study visit, the latter based on the type, frequency, and duration of recreational physical activity (Manson et al. 2002). When physical activity information was unavailable, it was defined as the value given at the last visit or the weighted mean between visits if data were available. Geocoded participant address-specific neighborhood socioeconomic status was a sum of Z-transformed U.S. Census tract-level measures of median household income; percent of households with interest dividends or rent income; percent of population at least 25 years of age with a high school degree; percent of population at least 16 years of age with professional, managerial, or executive occupations; and median value of owner-occupied housing units (Diez Roux et al. 2001). Geocoded participant address-specific daily mean temperature, dew point, and barometric pressure were averaged across all National Climatic Data Center monitoring stations within 50 km (NCDC 2019), then averaged over 2, 7, 28, and 365 d prior to and including the day of the study visit.

Subpopulation-specific covariates included sex (in ARIC), randomly assigned treatment group (in WHI), case–control status (in WHI-AS311 and WHI-BAA23), and other sampling-related variables in WHI-AS311 (i.e., enrollment year, age at enrollment, follow-up time, DNA extraction method).

Exclusions
Of all observations in WHI and ARIC (n = 285,548), small percentages were excluded because they were made on participants in one WHI center outside of the contiguous 48 states (2%), on study visit dates for which PM was not estimable (2%), among participants with a study-specific leukocyte count <99.5th percentile (leukopenia, 0.5%), or conditions associated with abnormal leukocyte traits such as hematological malignancy (1.7%) or oral/parenteral use of a granulocyte/macrophage colony stimulating factor (<0.01%), lithium (0.2%), glucocorticosteroid (1.1%), or antibiotic use (2.6%).

Multiple Imputation
To avoid potential for selection bias in complete-data analyses when data are missing at random (Hernán et al. 2004), multivariate imputation by chained equations (MICE) (Azur et al. 2011; Stuart et al. 2009) was used to impute missing data (percentage missing range: 0.6–9.1%). Binary and categorical data were imputed using logistic regression, and continuous variables were imputed using predictive means matching.

Attrition Weights
To address the potential for bias due to nonrandom attrition over time in leukocyte count analyses in WHI and ARIC, stabilized inverse probability weights for each participant were calculated at each examination using logistic regression, where the numerator was the marginal probability of the participant not being lost to follow-up at an examination and the denominator was the probability of the participant not being lost to follow-up at an examination conditional on their covariate patterns at prior examination (Howe et al. 2016).

Statistical Analysis: Leukocyte Count
Study- and center-stratified, PM–leukocyte count associations were estimated using an attrition-weighted and covariate-adjusted, two-level, linear, mixed-effects, longitudinal model including a random intercept for examination at the participant level. The model was given by

\[ LC_{ij} = \beta_0 + \beta_1 PM_{ij} + \beta_2 Z_{ij} + b_{0i,j} + e_{ij}, \]

where \( i \) and \( j \) denote the \( i \)-th examination (level 1) of the \( j \)-th participant (level 2); \( LC \) is the leukocyte count; \( \beta_0 \) is the intercept; \( PM \) is the 2- or 7-d mean of PM10 or the 1- or 12-month mean of PM2.5, PM10, or PM2.5–10; and \( Z \) is a vector of covariates. The term \( \left( b_{0i,j} \right) \sim N(0,G) \) is a random intercept for examination at the participant level to account for within-participant variation, and \( e_{ij} \sim N(0,\sigma^2) \) is the random error at the examination level. Study- and center-specific measures of association (\( \beta_1 \)) and their 95% confidence intervals (CIs) were estimated as \( \beta_1 \pm 1.96 \times \text{the standard error (SE)} \) per 10-\( \mu \)m\(^2\) increase in PM, forest plotted, and pooled in random-effects meta-analyses (DerSimonian and Laird 1986) after testing homogeneity of associations among strata (\( p_{\text{Cochran}} > 0.10 \)) (Cochran 1954).

Statistical Analysis: Leukocyte Composition
Subpopulation-stratified, cross-sectional, PM–leukocyte proportion associations were analyzed using multivariate methods for compositional data (Aitchison 1982; Egozcue et al. 2003), that is, a set of positive, mutually exclusive components (such as proportions, \( p \)) that represent parts constituting a whole, are multicollinear, and collectively sum to 1 within a constrained space called a simplex. Proportions were isometrically log-ratio (ilr)-transformed from the simplex to real (Euclidean geometric) space. Transformation—which allowed for the dependent variation (Chastin et al. 2015; Egozcue et al. 2003) and relative positioning of components in the simplex (Chastin et al. 2015; Fairclough et al. 2017)—resulted in \( p \)-1 orthogonal (i.e., non-multicollinear) coordinates. It also allowed for back-transformation of multivariate results into component proportions (Pawlowsky-Glahn et al. 2017).
Back-transformation was based on compositional data analysis models, as given by

$$ilr(LP) = \beta_0 + \beta_1 PM + \beta_2 Z + \epsilon,$$  \[2\]

where $ilr(LP)$ denotes the ilr-transformed estimated leukocyte proportions; $\beta_0$ is the intercept; $PM$ is the 2- or 7-d mean of PM$_{10}$ or the 1- or 12-month mean of PM$_{2.5}$, PM$_{10}$, or PM$_{2.5-10}$; $Z$ is a vector of covariates; and $\epsilon \sim (0, \sigma^2)$ is the random error term. The vector of association measures ($\beta_1$) denotes the five orthogonal coordinates, the back-transformation of which represents the corresponding difference in each of the six leukocyte proportions per 10-$\mu$g/m$^3$ increase in PM. Because the SEs of $\beta_1$ cannot be back-transformed, the SEs of back-transformed leukocyte proportion associations were estimated using 1,000 bootstrap samples. Subpopulation-specific measures of association were reported as absolute percentage differences ($\%$), forest plotted, and pooled in random effects meta-analyses (DerSimonian and Laird 1986) after testing homogeneity of associations among strata ($p_{Cochran}$’s $Q < 0.10$) (Cochran 1954).

**Statistical Analysis: Sensitivity**

In leukocyte count analyses, Model 1 adjusted for self-identified race/ethnicity, age, sex (in ARIC), randomly assigned treatment group (in WHI), visit, mean temperature, mean dew point, mean barometric pressure, season (to control for within-year variation), and a restricted cubic natural spline function of calendar date (Bhaskaran et al. 2013; Dominici et al. 2002; Peng et al. 2006) with one knot per year to control for secular trends in PM and leukocyte count methods. Model 2 also adjusted for potential socioeconomic confounders (individual-level education and neighborhood socioeconomic status). Model 3 additionally adjusted for behavioral variables that explain variation in leukocyte traits or account for residual confounding (smoking status, alcohol use, BMI, and physical activity). The sensitivity of Model 3 results to the use of two knots per calendar year, one knot for every two calendar years, and no calendar date adjustment was assessed. Although leukocyte composition analyses also adjusted for subpopulation-specific covariates, the models did not adjust for calendar date because leukocyte proportions were estimated using the same methods across subpopulations. In addition, leukocyte composition models were not center-stratified due to small sample sizes and instead were adjusted for U.S. Census region (Midwest, Northeast, South, and West). Sensitivity of leukocyte count associations to PM estimation method was examined by substituting spatially estimated 28- and 365-d mean concentrations of PM$_{10}$ for spatiotemporally estimated 1- and 12-month mean concentrations of PM$_{10}$. Sensitivity of significant PM-estimated leukocyte composition associations were assessed in a subset of ARIC participants with available measured leukocyte composition data (lymphocyte, monocyte, and granulocyte proportions). Additional sensitivity of PM–leukocyte composition associations were evaluated by estimating PM associations with the log-transformed ratio of CD4$^+$ to CD8$^+$ T-cell proportions (CD4:CD8)—a marker of immune function and

Figure 1. Map of geocoded Women’s Health Initiative (1993–2002) and Atherosclerosis Risk in Communities (1986–1998) study participants and centers at baseline. WHI centers ($n = 39$) followed 1,238–3,690 participants. ARIC centers followed 3,588–3,943 participants. WHI and ARIC centers were co-located in Minneapolis, MN, and Winston-Salem, NC.
Table 2. Mean ± standard deviation (SD) particulate matter concentrations among n = 165,675 participants with leukocyte count data before imputation, Women’s Health Initiative (1993–2002) and Atherosclerosis Risk in Communities (1986–1998) study.

<table>
<thead>
<tr>
<th></th>
<th>WHI screening visit and ARIC visit 1</th>
<th>WHI Annual visit 3</th>
<th>ARIC</th>
<th>WHI and ARIC Percentage imputed of 285,548 observations (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM (µg/m³)</td>
<td>n = 159,162</td>
<td>n = 144,744</td>
<td>n = 77,096</td>
<td></td>
</tr>
<tr>
<td>PM₁₀</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-d</td>
<td>29.5 ± 11.9</td>
<td>28.4 ± 11.1</td>
<td>28.4 ± 11.2</td>
<td>39.8 ± 14.1</td>
</tr>
<tr>
<td>7-d</td>
<td>28.7 ± 9.3</td>
<td>27.6 ± 8.3</td>
<td>27.6 ± 8.6</td>
<td>39.2 ± 10.3</td>
</tr>
<tr>
<td>1-month</td>
<td>20.9 ± 6.7</td>
<td>20.6 ± 6.6</td>
<td>20.6 ± 6.6</td>
<td>25.2 ± 7.1</td>
</tr>
<tr>
<td>12-month</td>
<td>20.9 ± 5.1</td>
<td>20.8 ± 5.1</td>
<td>20.7 ± 5.0</td>
<td>24.4 ± 4.4</td>
</tr>
<tr>
<td>PM₂·₅</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-month</td>
<td>12.2 ± 4.3</td>
<td>12.0 ± 4.1</td>
<td>12.0 ± 4.2</td>
<td>15.2 ± 5.2</td>
</tr>
<tr>
<td>12-month</td>
<td>12.1 ± 3.0</td>
<td>12.0 ± 3.0</td>
<td>12.0 ± 2.9</td>
<td>14.7 ± 3.6</td>
</tr>
<tr>
<td>PM₂·₅·₁₀</td>
<td>1-month</td>
<td>8.7 ± 4.7</td>
<td>8.6 ± 4.8</td>
<td>8.6 ± 4.8</td>
</tr>
<tr>
<td>12-month</td>
<td>8.7 ± 3.9</td>
<td>8.7 ± 4.0</td>
<td>8.7 ± 4.0</td>
<td>9.7 ± 2.1</td>
</tr>
</tbody>
</table>

Note: ARIC, Atherosclerosis Risk in Communities; CI, confidence interval; PM, particulate matter; PM₁₀, PM ≤ 10 µm in diameter; PM₂·₅, PM ≤ 2.5 µm in diameter; PM₂·₅·₁₀, PM > 2.5 and < 10 µm in diameter; WHI, Women’s Health Initiative.
*WHI Observational Study participants only.
†Participants from Forsyth County only.
‡Participants from Washington County only.

Results
Of the 150,328 WHI participants and 15,347 ARIC participants with leukocyte count data (total n = 165,675; Figure 1), 96% and 94% had baseline data after exclusions. At baseline, participants were 62.3 years of age on average and mostly female (96%), white (84%), more than half had high school educated (74%), never/former smokers (91%), and current alcohol users (70%). Mean BMI, physical activity, and leukocyte count were 28.0 kg/m², 12.3 MET-hours/week, and 5,908 cells/µL (Table 1). Participants in the WHI and ARIC subpopulations with leukocyte composition data (n = 8,457; Table S2) were more likely to be younger (mean age: 61.5 y) and male (16%) and less likely to be white (45%), more than half had high school educated (52%), never/former smokers (85%), and current alcohol users (52%) than those with leukocyte count data. Among these subpopulations, mean estimated leukocyte cell type percentages were 9% (CD8⁺ T cells), 18% (CD4⁺ T cells), 7% (natural killer cells), 7% (B cells), 10% (monocytes), and 49% (granulocytes).

Possible biomarker for coronary heart disease (Neupane et al. 2019). PM–CD4:CD8 associations were reported as percentage changes.

Table 3. Pooled difference in leukocyte count (Δ, cells/µL) per 10-µg/m³ increase in PM concentrations among n = 165,675 participants, Women’s Health Initiative (1993–2002) and Atherosclerosis Risk in Communities (1986–1998) study.

<table>
<thead>
<tr>
<th>PM exposure (µg/m³)</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Δ (95% CI) cells/µL</td>
<td>pCochran’s Q</td>
<td>Δ (95% CI) cells/µL</td>
</tr>
<tr>
<td>PM₁₀</td>
<td>-6 (-12, 0)</td>
<td>0.89</td>
<td>-7 (-12, -1)</td>
</tr>
<tr>
<td>2-d mean</td>
<td>-10 (-19, -1)</td>
<td>0.49</td>
<td>-10 (-20, -1)</td>
</tr>
<tr>
<td>7-d mean</td>
<td>22 (3, 41)</td>
<td>2.5 × 10⁻³</td>
<td>8 (-8, 25)</td>
</tr>
<tr>
<td>1-month mean</td>
<td>65 (26, 103)</td>
<td>6.5 × 10⁻⁴</td>
<td>32 (4, 59)</td>
</tr>
<tr>
<td>PM₂·₅ (µg/m³)</td>
<td>33 (9, 56)</td>
<td>0.21</td>
<td>21 (0, 43)</td>
</tr>
<tr>
<td>1-month mean</td>
<td>144 (65, 163)</td>
<td>0.59</td>
<td>64 (15, 114)</td>
</tr>
<tr>
<td>PM₂·₅·₁₀ (µg/m³)</td>
<td>18 (8, 44)</td>
<td>0.01</td>
<td>-1 (-24, 21)</td>
</tr>
<tr>
<td>1-month mean</td>
<td>67 (8, 127)</td>
<td>6.5 × 10⁻⁶</td>
<td>18 (-30, 66)</td>
</tr>
</tbody>
</table>

Note: ARIC, Atherosclerosis Risk in Communities; CI, confidence interval; PM, particulate matter; PM₁₀, PM ≤ 10 µm in diameter; PM₂·₅, PM ≤ 2.5 µm in diameter; PM₂·₅·₁₀, PM > 2.5 and < 10 µm in diameter; WHI, Women’s Health Initiative.
*Model 1 adjusted for race/ethnicity, age, sex (in ARIC), randomly assigned treatment group (in WHI), mean temperature, mean dew point, mean barometric pressure, season, and a restricted cubic natural spline function of calendar time with one knot per calendar year.
†Model 2 adjusted for all covariates in Model 1 and additionally for individual-level education and neighborhood socioeconomic status.
‡Model 3 adjusted for all covariates in Model 2 and additionally for smoking status, alcohol use, body mass index, and physical activity.
§Homogeneity of associations among strata was tested using Cochran’s Q-test statistic, where a pCochran’s Q < 0.10 suggests there is evidence to reject the null hypothesis of homogeneity.
the leukocyte count (i.e., they had wide CIs). However, the associations also were attenuated by additional adjustment for potential socioeconomic confounders (Model 2) and behavioral variables (Model 3). For example, there were 114 (65, 163), 64 (15, 114), and 28 (−20, 75) cells/µL higher leukocyte counts per 10-µg/m³ increase in the 12-month mean PM2.5 concentration in Models 1–3 (Table 3; Figure S2). In sensitivity analyses, estimates were generally robust to variation in the method of controlling for calendar date (see Figure S3). Leukocyte count associations with 28- and 365-d mean PM10 concentrations also were imprecise and no different from the null associations (see Table S4), as those between leukocyte count and 1- and 12-month mean PM10.

Across PM size fractions and averaging durations, PM–leukocyte compositional associations in Model 3 (Table 4) differed little from those in Models 1 and 2 (see Tables S5–S6). Higher 7-d mean PM10 concentrations were associated with somewhat higher CD8⁺ T-cell proportions, whereas 1- and 12-month mean PM10 concentrations were associated with somewhat lower CD8⁺ T-cell proportions (Table 4; Figure S4). One- and 12-month mean concentrations of PM2.5 were associated with lower CD8⁺ T-cell, NK cell, and B-cell proportions and higher granulocyte proportions. For example, there was a 1.1% (−1.9%, −0.3%) lower CD8⁺ T-cell proportion and 1.2% (0.6%, 1.8%) higher granulocyte proportion per 10-µg/m³ increase in 1-month mean PM2.5 (Figure 3). In contrast, there were 0.6% (−1.3%, 0.1%) and 1.2% (−2.4%, 0.1%) lower granulocyte proportions per 1- and 12-month mean PM2.5–10 (see Figure S4). PM2.5 associations with estimated granulocyte proportions were consistent in magnitude and direction with those in the analyses of measured granulocyte proportions (see Table S7). PM–CD4:CD8 associations were generally inconsistent, with suggestively inverse associations with short-duration PM10 and suggestively positive associations with longer duration PM10 and PM2.5; however, CIs were wide and included the null (Table S8).

Discussion

Results from this study suggest that mid- to longer-duration exposures to PM2.5 concentrations below U.S. EPA NAAQS may be associated with a higher leukocyte count, higher granulocyte proportion, and lower CD8⁺ T-cell proportion among multi-ethnic and geographically diverse populations of U.S. women and men. Although leukocyte count associations were also observed with 1- and 12-month mean PM10 and PM2.5–10 concentrations, adjusting for potential socioeconomic confounders attenuated them. Indeed, lower socioeconomic status has been related both to increases in CVD risk (Elo 2009) and higher concentrations of ambient PM (Hajat et al. 2015). Further attenuation was observed with additional adjustment for behavioral variables (smoking, alcohol use, BMI, and physical activity) suggesting that they may account for residual confounding by socioeconomic or other unmeasured characteristics. Taken together with prior evidence suggesting positive (Chen and Schwartz 2008) and null (Viehmann et al. 2015) associations between longer-duration PM10 with leukocyte counts, the present results were unable to clarify the relationship. Nevertheless, positive—yet imprecise—leukocyte count estimates remained for PM2.5, supporting evidence first reported in the Heinz Nixdorf Recall study (Viehmann et al. 2015). Moreover, the magnitudes of estimates presently observed are on par with those previously associated with a 1-cigarette/d increase in smoking (Hansen et al. 1990; Pettiti and Kipp 1986; Schwartz and Weiss 1991).

PM2.5 concentrations were also associated with leukocyte composition; particularly, with higher granulocyte and lower CD8⁺ T-cell proportions. This observation is consistent with results from the Social Environment and Biomarkers of Aging Study (SEBAS) in Taiwan that found positive associations between long-duration PM2.5 exposure and the proportion of neutrophils, the most abundant type of granulocyte (Chuang et al. 2011). SEBAS also detected similar associations with long-duration PM10 concentrations, but...
Persistent systemic inflammation due to longer-duration PM exposure is a biologically plausible mechanism linking PM with adverse health. Indeed, systemic inflammation has been implicated in endothelial injury, atherosclerotic disease progression, and subsequent increases in CVD risk (Ross 1999). In the epidemiologic context, systemic inflammation, as measured by leukocyte count, has been consistently and independently associated with CVD and mortality in WHI (Kabat et al. 2017; Margolis et al. 2005), in ARIC (Lee et al. 2001), and in other populations (Brown et al. 2001; Danesh et al. 1998; Ruggiero et al. 2007).

The results presented herein support the hypothesis that chronic exposure to PM contributes to systemic inflammation and may partly explain the established connection between PM and CVD risk (Chi et al. 2016a; Miller et al. 2007). They support prior studies that mechanistically linked atherosclerosis and the inflammatory responses to PM (Adar et al. 2013; Brook and Rajagopalan 2010; Diez Roux et al. 2008; Künzli et al. 2005; Perez et al. 2015). Such studies observed higher pro-inflammatory cytokines following inhalation and deposition of PM in the lungs (Pope et al. 2016; Tan et al. 2000; Terashima et al. 1997a, 1997b; van Eeden et al. 2001) and the activation of coagulation and adhesion molecules (Baccarelli et al. 2007; Bind et al. 2012; O’Neill et al. 2007; Pope et al. 2016; Rückerl et al. 2006; Tsai et al. 2012), which could ultimately lead to increased leukocyte content within and vulnerability to rupture of atherosclerotic plaques (Brook and Rajagopalan 2010; Madjid et al. 2004; Ross 1999).

Although the inverse relationship between short-duration (i.e., 2- and 7-d mean) ambient PM$_{10}$ exposures and leukocyte counts may be at odds with this suggestion, PM exposure may initiate pulmonary alveolar microvascular sequestration of monocytes and granulocytes (Goto et al. 2004; Terashima et al. 1999; Yatera et al. 2008), thereby reducing their concentrations in peripheral blood over the short term (Ghio et al. 2003; Yatera et al. 2008). Animal studies of monocytes and acute PM$_{10}$ exposure also suggest that atherosclerotic plaques may recruit leukocytes from the circulation (Yatera et al. 2008). The inverse PM$_{10}$–leukocyte count associations with short-duration exposure in the present study are in contrast to null (Liao et al. 2005; Seaton et al. 1999) and positive (Schwartz 2001; Steinvil et al. 2008) epidemiologic associations observed in other contexts. However, they are consistent with observed inverse associations with short-duration exposure to PM$_{2.5}$ in the Normative Aging Study (Zeka et al. 2006).

The characterization of PM–leukocyte associations in the compositional context is particularly relevant given the increasing availability of epigenomic biomarkers that are based on DNA extracted from peripheral blood with leukocyte proportions that can vary widely among participants. However, leukocyte cell types possess distinguishing patterns of DNA methylation that are driven in part by leukocyte composition (Jaffe and Irizarry 2014). Common practice is therefore to restrict measurement of DNA methylation to a single cell type (Chi et al. 2016b), to statistically adjust associations with DNA methylation to leukocyte proportions determined via cytomtery as part of a complete blood count/differential, or in its...
Indeed, leukocyte composition may itself be a mediator of PM2.5 associations. As such, DNAm associations with PM2.5 mechanisms that involve inflammatory responses among racially, ethnically, and environmentally diverse U.S. populations in U.S. EPA Regions 1–10.

The present results are nevertheless limited by the variances of the observed association estimates. The analyses were weighted for attrition to avoid potential selection bias due to nonrandom loss to follow-up; however, the loss of bias came at the cost of precision (Cole and Hernán 2008). Furthermore, precision was influenced by technical, temporal, and biological variation of leukocyte count measurements. Participant blood samples were collected, processed, and analyzed by local laboratories across the United States using different automated hematology cell counters. Indeed, secular trends in methods of determining leukocyte count (Ruggiero et al. 2007) may have affected the precision or accuracy of association estimates. And while lack of adjustment for other cell (e.g., erythrocyte, platelet) counts capable of explaining some variation in leukocyte counts may have contributed to the precision of estimates observed herein, there also is evidence to suggest high within-laboratory reliability of leukocyte counts (Nieto et al. 1992) and robustness of study- and center-stratified, longitudinal model results to multiple methods of calendar date adjustment. Moreover, erythrocyte and platelet counts—plausible intermediates of PM-leukocyte count associations—were neither uniformly available nor necessarily appropriate candidates for statistical adjustment (Schisterman et al. 2009).

Additional limitations include error in estimated leukocyte proportions and PM concentrations. Although cytometrically determined leukocyte proportions for the cell types of interest were not available herein at participant visits with corresponding PM data, estimation of the CD8+ T-cell, CD4+ T-cell, NK cell, B-cell, monocyte, and granulocyte proportions at hand was associated with a low root mean square error (median rMSE: 8.2%, range: 5.4–11.6%) (Houseman et al. 2012; Koestler et al. 2013). Furthermore, the validity of spatially estimated daily PM10 estimates was demonstrated with an average prediction error and standardized prediction error near zero, a root mean square standardized near one, and a root mean square prediction error near the SE (Liao et al. 2006, 2007). Similarly, models for spatiotemporally estimated monthly mean PM10 and PM2.5 estimation performed well, with high squared Pearson correlations between excluded monthly observations and model predictions ($R^2 = 0.68–0.77$) in a 5- to 10-fold, out-of-sample cross-validation (Yanosky et al. 2014). Therefore, outcome and exposure measurement error were less likely to have biased observed associations.

Limitations aside, this longitudinal study observed that 1- and 12-month mean ambient PM2.5 concentrations were associated with higher leukocyte count. It is the first to do so in large, multiethnic and geographically diverse populations of women and men from two well-characterized cardiovascular disease cohorts. Furthermore, this study is the first to use compositional data analysis methods to estimate associations between ambient PM2.5 concentrations and leukocyte composition. Its analyses accounted for known relationships among proportions, thereby avoiding methodological biases inherent in conventional analyses that erroneously assume compositional data are independent. Results from them are therefore relatively well positioned to inform future causal analyses using leukocyte-derived biomarkers.

In conclusion, findings suggest that mid- to longer-duration ambient exposure to fine particulate matter (PM2.5) air pollution may induce subclinical, but epidemiologically important, inflammatory responses among racially, ethnically, and environmentally diverse U.S. populations in U.S. EPA Regions 1–10.

**Acknowledgments**

The authors thank the staff and participants of the ARIC study for their important contributions. The Atherosclerosis Risk in Communities (ARIC) study was supported by National Heart, Lung, and Blood Institute (NHLBI) grants 5U01-HL080963, 5U01-HL080967, 5U01-HL080968, 5U01-HL080970, and 5U01-HL080973 and National Institute of Environmental Health Sciences grants U01ES023724 and U01ES023725. The ARIC Study Investigators and Coordinators also thank the atherosclerosis risk in communities (ARIC) study for their important contributions. The Atherosclerosis Risk in Communities (ARIC) study was supported by National Heart, Lung, and Blood Institute (NHLBI) grants 5U01-HL080963, 5U01-HL080967, 5U01-HL080968, 5U01-HL080970, and 5U01-HL080973 and National Institute of Environmental Health Sciences grants U01ES023724 and U01ES023725. The ARIC Study Investigators and Coordinators also thank the atherosclerosis risk in communities (ARIC) study for their important contributions.
Risk in Communities study has been funded in whole or in part with Federal funds from the Department of Health and Human Services (DHHS), National Institutes of Health (NIH), National Heart, Lung, and Blood Institute (NHLBI) (contracts HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I, and HHSN268201700005I). Funding was also supported by NHLBI through the American Recovery and Reinvestment Act of 2009 (ARRA) 5RC2HL102419 and National Institute of Neurological Disorders and Stroke R01NS087541. Data from the ARIC study are available on request at https://www2.cscce.unc.edu/aric/distribution-agreements.

The WHI program is funded by the NHLBI through contracts HHSN268201100046C, HHSN268201100041C, HHSN268201100026C, HHSN268201100036C, and HHSN271201100004C. WHI-AS311 was supported by American Cancer Society award 125,299-RSG-13-100-01-CCE. and HHSN268201100046C. Data from the WHI are available on request at https://www2.cscc.unc.edu/aric/distribution-agreements.

References


