

1 Environmental Exposure to Volatile Organic Compounds is Associated with Endothelial
2 Injury

3

4

5 *Daniel W. Riggs,¹⁻⁵ *Marina V. Malovichko,¹⁻⁴ Hong Gao,¹⁻⁴ Katlyn E. McGraw,¹⁻⁴
6 Breandon S. Taylor,^{1,2} Tatiana Krivokhizhina,¹⁻⁴ Shesh N. Rai,^{1,2,6} Rachel J. Keith,^{1,2,4}
7 Aruni Bhatnagar,¹⁻⁴ and Sanjay Srivastava¹⁻⁴

8

9 ¹Superfund Research Center, University of Louisville; ²Diabetes and Obesity Center,
10 University of Louisville; ³Christina Lee Brown Envirome Institute, University of Louisville,
11 ⁴Division of Environmental Medicine, Department of Medicine, University of Louisville;
12 ⁵Department of Epidemiology and Population Health, University of Louisville;
13 ⁶Department of Bioinformatics and Biostatistics, University of Louisville

14

15 *These people contributed equally to the manuscript

16

17

18

19 **Address correspondence to:** Sanjay Srivastava, PhD
20 Superfund Research Center, Cardiovascular
21 Innovation Institute
22 302 E. Muhammad Ali Blvd
23 Louisville, KY 40202
24 Phone: 502-852-5724; Fax: 502-852-5834
25 E-mail: sanjay.srivastava@louisville.edu

26

27

28

29

30

31

32

33

34

35

36

37 **ABSTRACT**

38 **Objective:** Volatile organic compounds (VOCs) are airborne toxicants abundant in
39 outdoor and indoor air. High levels of VOCs are also present at various Superfund and
40 other hazardous waste sites; however, little is known about the cardiovascular effects of
41 VOCs. We hypothesized that ambient exposure to VOCs exacerbate cardiovascular
42 disease (CVD) risk by depleting circulating angiogenic cells (CACs).

43 **Approach and Results:** In this cross-sectional study, we recruited 603 participants with
44 low-to-high CVD risk and measured 15 subpopulations of CACs by flow cytometry and
45 16 urinary metabolites of 12 VOCs by LC/MS/MS. Associations between CAC and VOC
46 metabolite levels were examined using generalized linear models in the total sample, and
47 separately in non-smokers. In single pollutant models, metabolites of
48 ethylbenzene/styrene and xylene, were negatively associated with CAC levels in both the
49 total sample, and in non-smokers. The metabolite of acrylonitrile was negatively
50 associated with CD45^{dim}/CD146⁺/CD34⁺/AC133⁺ cells and CD45⁺/CD146⁺/AC133⁺, and
51 the toluene metabolite with AC133⁺ cells. In analysis of non-smokers (n=375),
52 multipollutant models showed a negative association with metabolites of
53 ethylbenzene/styrene, benzene, and xylene with CD45^{dim}/CD146⁺/CD34⁺ cells,
54 independent of other VOC metabolite levels. Cumulative VOC risk score showed a strong
55 negative association with CD45^{dim}/CD146⁺/CD34⁺ cells, suggesting that total VOC
56 exposure has a cumulative effect on pro-angiogenic cells. We found a non-linear
57 relationship for benzene, which showed an increase in CAC levels at low, but depletion
58 at higher levels of exposure. Sex and race, hypertension, and diabetes significantly
59 modified VOC associated CAC depletion.

60 **Conclusion:** Low-level ambient exposure to VOCs is associated with CAC depletion,
61 which could compromise endothelial repair and angiogenesis, and exacerbate CVD risk.

62 1. INTRODUCTION

63 Exposure to ambient air pollution has been widely recognized to be an important
64 contributor to premature mortality and morbidity. The Global Burden of Disease (GBD)
65 study estimates that in 2015, 4.2 million deaths could be attributed to ambient air
66 pollutants, with an additional 2.8 million deaths due to household air pollution.¹ Exposure
67 to fine particulate matter in the ambient air, defined as particle size < 2.5 µm in diameter
68 (PM_{2.5}), is associated with an increased risk for hypertension,² myocardial infarction,³
69 type 2 diabetes,^{4,5} stroke,⁶ and cardiovascular mortality.⁷ The emphasis on PM_{2.5} and
70 cardiovascular disease (CVD) has been motivated largely by the availability of historical
71 monitoring data, however in an urban setting, over 90% of pollutant mass consists of
72 gases or vaporous compounds, such as volatile organic compounds (VOCs).⁸ In
73 comparison with PM_{2.5}, data on VOCs tend to exhibit higher temporal and spatial
74 variability.

75 Most studies on VOC have focused on tobacco smoke⁹⁻¹¹ which represents a
76 major source of human exposures. However, non-tobacco derived VOCs are ubiquitous
77 in urban environments. High concentrations of VOCs such as alkanes, alkenes, aromatic
78 hydrocarbons, and aldehydes, are emitted into the atmosphere from both anthropogenic
79 and biogenic sources, and although biogenic VOCs are more abundant worldwide, in
80 urban areas, anthropogenic VOCs are often higher.¹² Outdoor VOC exposures can
81 originate from vehicle emissions, industrial facilities, landfills and hazardous waste sites,
82 while indoor exposure can originate from diverse sources, such as paints, cleaning

83 agents, personal care products, and solvents.¹³ Our analysis of the secular trends of
84 VOC exposures in the US suggests that even though there was a decline in the ambient
85 air VOC levels, individual-level exposure increased from 2005-2014.¹⁴ Hence, there is
86 urgent need to assess the health impacts of VOC exposures.

87 VOC exposure have been associated with a wide range of acute and chronic
88 health effects, including allergies, respiratory diseases, liver dysfunction, and cancer.¹⁵,
89 ¹⁶ Moreover, exposures to ambient levels of acrolein, benzene, and xylene have been
90 associated with elevated CVD risk.^{17, 18} In an intra-urban analysis of VOC gradients,
91 higher levels of VOCs were found to be associated with increased cardiovascular and
92 cancer mortality.¹⁶ Moreover, longitudinal time-series studies have shown that ambient
93 VOCs, such as alkenes and alkynes, are associated with increased heart failure
94 emergency hospitalizations,¹⁹ as well as overall cardiovascular hospitalizations,²⁰ and
95 that occupational exposure to VOC such as styrene increases the relative risk of death
96 from ischemic heart disease.²¹

97 The biological plausibility that VOCs can induce cardiovascular injury and
98 exacerbate CVD is extensively supported by a range of animal studies showing that
99 exposures to benzene, acrolein, and formaldehyde, can lead to changes in blood
100 pressure regulation,²²⁻²⁴ depletion of circulating stem cells,²⁵ insulin resistance,²⁶ and
101 platelet activation.²⁷ Exposure to these VOCs also induces systemic dyslipidemia,
102 lipoprotein modification,²⁸ dilated cardiomyopathy;²⁹ and lead to the acceleration of
103 atherogenesis³⁰ and exacerbation of myocardial ischemic injury.³¹ Nonetheless, despite
104 this evidence, little is known about the effects of VOCs on the mechanisms and processes
105 that contribute to cardiovascular disease in humans. Previous work has shown that in

106 humans, high levels of VOC exposures from sources such as cigarette smoke^{32, 33} and
107 diesel exhaust³⁴ are associated with endothelial dysfunction. We have previously found
108 that exposure to acrolein and 1,3-butadiene are associated with lower reactive hyperemia
109 index, a measure of endothelial function.^{13, 35} The endothelium is a first line of defense
110 of cardiovascular tissues against environmental toxins, and endothelial function is a
111 sensitive reflection of vascular and systemic health. Therefore, to identify and understand
112 the effects of VOC exposures and cardiovascular health, we measured circulating
113 angiogenic cells (CAC), also referred to as endothelial progenitor cells (EPC). It is
114 believed that these cell aid in the repair of the endothelium after injury, and are good
115 biomarkers of vascular function.^{36, 37} CAC levels are associated with several CVD risk
116 factors,³⁸ and are predictive of both cardiovascular events and cardiovascular mortality.
117³⁹ Thus, using CACs as surrogate biomarkers of endothelial health, we examined their
118 association with urinary levels of VOC metabolite levels in participants with low to severe
119 CVD risk.

120 **2. METHODS**

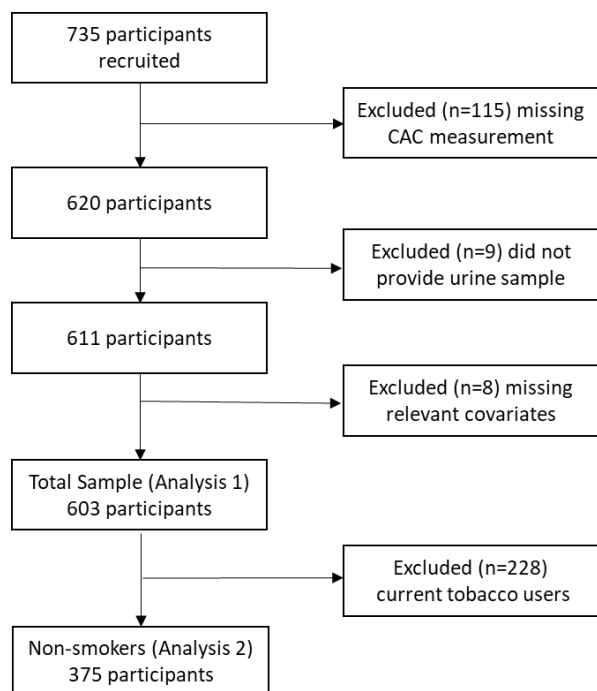
121 **2.1 Study Population and Design**

122 In a cross-sectional study, adult study participants, aged 25-70 years, were
123 recruited from the Louisville, KY area. Recruitment occurred in the summer months (May
124 to September) in the years 2018 and 2019, to minimize seasonal outdoor VOC
125 heterogeneity. Persons unwilling or unable to provide informed consent or with significant
126 and/or severe comorbidities were excluded. These included subjects that had hepatitis,
127 HIV/AIDS, active treatment for cancer, and active bleeding including wounds. Other
128 exclusions included subjects with body weight less than 100 pounds, pregnant women,

129 prisoners, and other vulnerable populations, those with conditions known to effect
130 peripheral blood cell counts and bone marrow function.

131 At the time of enrollment non-fasting blood and spot urine were collected from
132 patients who met the enrollment criteria and gave written consent. Demographic
133 information, including age, sex, ethnicity, residential address, smoking status history and
134 secondhand smoke exposure (verified by measurement of urinary cotinine
135 concentrations), and BMI (calculated from self-reported height and weight) was obtained
136 by using a questionnaire. Information on medication usage, blood pressure, and CVD
137 history, including incidence of heart attack, heart failure, angina, hypertension,
138 hypercholesterolemia, diabetes, obesity, stroke, and revascularization were also obtained
139 through the questionnaire. For the current study, hypertension was defined as a previous
140 diagnosis of hypertension or systolic blood pressure greater than 140 mmHg. Diabetes
141 was defined as previous diagnosis of diabetes or HbA1C greater than 6.5. The study was
142 approved by the University of Louisville Institutional Review Board (IRB 15.1260).

143 A total of 735 individuals met the inclusion criteria (Scheme 1).



Scheme 1. Recruitment flow chart.

144

145 Of those enrolled in the study, 115 individuals did not provide blood samples, or did not
146 have circulating angiogenic cell measurements, 9 individuals did not provide urine
147 samples, and 8 were missing other covariates. The final sample consisted of 603
148 participants with both urinary VOC metabolites and CAC measurements. For the non-
149 smoker analysis, participants (n=228) with urinary cotinine value >40 µg/g creatinine were
150 excluded, for a final sample size of n=375 participants.

151 2.2 Urinary Volatile Organic Compound Metabolites

152 Urinary metabolites of 12 VOCs (16 metabolites), and free forms of the nicotine
153 metabolite cotinine were quantified by ultra-performance liquid chromatography-mass
154 spectrometry (UPLC-MS/MS) as described before.^{9, 11, 40} All VOC metabolites that were
155 below the limit of detection (LOD) were imputed with the LOD value divided by the square
156 root of 2. All urinary analytes were normalized to urinary creatinine levels, measured on

157 a COBAS MIRA-plus analyzer (Roche, New Jersey) with Infinity Creatinine Reagent
158 (Thermo Fisher Scientific, Massachusetts). A complete list of metabolites and parent
159 compounds is provided in supplementary material (Table S1).

160 **2.3 Circulating Angiogenic Cells**

161 Fifteen different subtypes of circulating angiogenic cells were measured in the
162 peripheral blood by flow cytometry - on fresh blood specimens as previously described.
163 ^{17, 41} CACs were defined as CD14 negative cells and differentiated into subtypes based
164 on expression of CD45, CD34, CD146 and AC133 surface markers.

165 **2.4 Air Pollution and Meteorological Data**

166 Data on ambient levels of PM_{2.5} (particulate matter with an aerodynamic diameter
167 of $\leq 2.5 \mu\text{m}$; $\mu\text{g}/\text{m}^3$) and ozone (ppm) were obtained from three regional EPA-validated
168 monitoring stations in the Louisville, KY region. We used the average of the three monitors
169 for the day the study participants provided blood and urine samples.

170 Data on average daily temperature ($^{\circ}\text{F}$) and relative humidity (%) were obtained
171 from the National Oceanic and Atmospheric Administration's (NOAA) National Centers
172 for Environmental Information. Data was download for the Louisville International Airport
173 monitoring station, which is in close proximity to the study participants' residence, for the
174 day the study participants provided blood and urine sample. This data is publicly available
175 (<https://www.ncdc.noaa.gov/cdo-web/datatools/selectlocation>).

176 **2.5 Statistical Analysis**

177 Participant characteristics are expressed as mean \pm standard deviation (SD) for
178 continuous variables, and frequency (%) for categorical variables. Descriptive statistics

179 were conducted for all VOC metabolites. Because CACs were positive, and heavily right-
180 skewed, generalized linear models with the gamma distribution and log-link were used to
181 examine the association between CACs and VOC metabolites (Figure 1). Models used
182 natural log-transformed VOC metabolites and were adjusted based on prior literature
183 showing an association with VOC exposure or CAC levels.⁴¹⁻⁴⁷ Model adjustments for
184 full data set included: age, sex, race, BMI, diabetes status, hypertension status, daily
185 PM_{2.5} levels, and smoking status. There are a total of 240 unique comparisons between
186 CACs and VOC metabolites. To account for multiple testing, we applied the Benjamini-
187 Hochberg procedure to control the false-discovery rate (FDR).

188 **Analysis of non-smokers**

189 Models for non-smokers (urinary cotinine <40 µg/g creatinine) were adjusted for
190 age, sex, race, BMI, diabetes status, hypertension status, and daily PM_{2.5} levels. Pearson
191 correlations were performed across VOC metabolites and daily PM_{2.5}, ozone,
192 temperature, and relative humidity (Supplemental Figure 1). We examined potential non-
193 linear relationships between VOC metabolites and CACs using restricted cubic splines,
194 with tests for non-linearity using the likelihood ratio test. An environmental risk score
195 (ERS) representing cumulative VOC exposure risk was calculated for each CAC using a
196 modified version of the method developed by Park *et al.*⁴⁸ Briefly, (1) VOC metabolites
197 were log transformed, (2) each log-transformed VOC metabolite was regressed against
198 each CAC separately, while adjusting for covariates, (3) the environmental risk score for
199 each participant was calculated as the sum of log-transformed VOC metabolite x
200 (regression coefficient/standard error). If multiple VOC metabolites came from the same
201 parent compound, only one was used in the creation of the ERS. The final set of VOC

202 metabolites used to calculate the ERS consist of: 3HPMA, AAMA, CYMA, MU, MHBMA3,
203 HPMMA, AMCC, PGA, 2HPMA, MA, BMA, and 3MHA+4MHA. Multi-pollutant models
204 including all VOC metabolites that were independently associated with CACs were further
205 developed to test whether single VOC effects were confounded by correlated VOCs
206 (Pearson correlations provided in Supplemental Table 2). These models used log-
207 transformed metabolite values, and the same adjustments as single pollutant models.
208 Interactions between covariates and log VOC metabolites were tested in single pollutant
209 models. We stratified analyses by participant characteristics to examine whether effect
210 estimates differed by subgroups.

211 In addition to single-VOC models, we created models for multiple VOC exposures.
212 To assess multiple VOC exposures as a mixture, we applied a Bayesian Kernel Machine
213 Regression (BKMR) approach to flexibly model the relationship between the mixture of
214 VOC metabolites and each of the 15 CACs.⁴⁹ The BKMR is a semi-parametric technique
215 that models the exposure response relationship as a non-parametric kernel function of
216 the mixture components, adjusting for covariates parametrically. Implementation of
217 BKMR relied on a Gaussian kernel distribution, with flat priors. The Gaussian kernel
218 flexibly captures underlying functional forms without specifying the shape of the exposure-
219 response curve. This method can assess independent effects, estimate the overall
220 mixture, and assess for interactions among mixture components. We used R software
221 (version 4.0.2) and the bkmr package⁵⁰ to create posterior inclusion probabilities (PIPs)
222 that quantify the relative importance of each exposure in the model, assess univariate
223 and bivariate associations between mixture components and CACs, assess interactions

224 among VOC metabolites, and estimate the effect of the overall mixture on individual
225 CACs.

226 As a sensitivity analysis, we conducted generalized linear models additionally
227 adjusting one at a time for temperature or relative humidity, while adjusting for the
228 covariates as above. We then repeated the models adjusting for ozone in place of PM_{2.5},
229 while adjusting one at time for temperature, then relative humidity. All statistical analyses
230 were performed using SAS, version 9.4 (SAS Institute, Inc., Cary, North Carolina) and
231 GraphPad Prism, version 8 (GraphPad Software, La Jolla, California).

232 3. RESULTS

233 The demographic characteristics of study participants are summarized in Table 1.

234 Table 1. Participant Characteristics

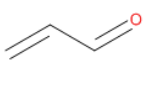
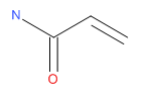
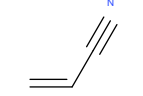
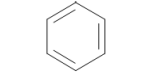
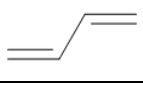
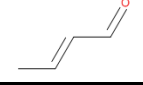
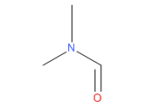
Variable	Total (n=603)	Non-Smokers (n=375)	Smokers (n=228)	p-value
Sex, Male	242 (40%)	135 (36%)	107 (47%)	0.008
Race				<0.001
White	464 (77%)	311 (83%)	153 (67%)	
Black	106 (18%)	43 (11%)	63 (28%)	
Other	33 (5%)	21 (6%)	12 (5%)	
Age (years)	49.7 ± 12.7	50.6 ± 13.2	48.3 ± 11.8	0.028
BMI (kg/m ²)	30.2 ± 6.8	30.7 ± 7.0	29.4 ± 6.4	0.023
SBP (mmHg)	127.9 ± 17.6	128.5 ± 18.1	127.1 ± 16.8	0.349
DBP (mmHg)	78.6 ± 11.4	78.1 ± 11.2	79.4 ± 11.8	0.178
Hypertension	247 (41%)	156 (42%)	91 (40%)	0.683
Diabetes	116 (19%)	65 (17%)	51 (22%)	0.128
MI	36 (6%)	24 (6%)	12 (5%)	0.632
Medication Use				
ASA	67 (12%)	46 (13%)	21 (10%)	0.295
Statin	82 (14%)	60 (17%)	22 (10%)	0.037
ACE or ARB	87 (15%)	63 (18%)	24 (11%)	0.047
Beta Blockers	53 (9%)	35 (10%)	18 (9%)	0.624

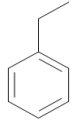
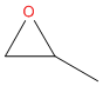
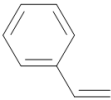
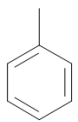
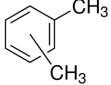
235 Values represent mean ± S.D. for continuous variables, and frequency (%) for categorical
236 variables. P-values are based on t-test for continuous variables, and Chi-squared test for

237 categorical variables. Abbreviations: BMI indicates body mass index; SBP, systolic blood
 238 pressure; DBP, diastolic blood pressure; MI, myocardial infarction; ASA, aspirin; ACE,
 239 angiotensin converting enzyme inhibitor; ARB, angiotensin receptor blocker.

240
 241 Compared with non-smokers, current smokers were more likely to be male (47% vs.
 242 36%), Black (28% vs. 11%), were of younger age, and have lower BMI. Although
 243 cardiovascular risk factors (blood pressure, hypertension, diabetes) were similar between
 244 the two groups, non-smokers were more likely to use statins, and ACE or ARB
 245 medications. Within the 375 non-smoking participants of the study, the average age was
 246 50.6 years. A majority were female (64%) and White (83%); 42% were hypertensive, and
 247 17% diabetic. The urinary metabolites of VOCs measured in study participants are listed
 248 in Table 2, along with their parent VOCs.

249 Table 2. Urinary levels of the major metabolites of volatile organic compounds (VOCs)
 250 in study participants.

Parent Structure	VOC	Metabolite	Total Mean \pm SD	Non-Smokers Mean \pm SD	n (%) LOD	Smokers Mean \pm SD	n (%) LOD
	Acrolein	CEMA	200.1 \pm 225.9	115.8 \pm 117.0	42 (11.2%)	338.6 \pm 286.0	7 (3.1%)
		3HPMA	879.3 \pm 1228.2	341.5 \pm 344.6	0 (0%)	1763.7 \pm 1594.1	0 (0%)
	Acrylamide	AAMA	109.9 \pm 120.3	72.4 \pm 61.6	1 (0.3%)	171.5 \pm 161.2	0 (0%)
	Acrylonitrile	CYMA	60.9 \pm 111.0	5.7 \pm 17.2	231 (61.6%)	151.8 \pm 137.4	7 (3.1%)
	Benzene	MU	146.0 \pm 195.7	133.4 \pm 214.0	6 (1.6%)	166.8 \pm 159.5	4 (1.8%)
	1,3-Butadiene	DHBMA	405.7 \pm 215.7	353.6 \pm 175.9	0 (0%)	491.2 \pm 246.2	0 (0%)
		MHBMA3	28.7 \pm 39.8	9.1 \pm 8.2	18 (4.8%)	60.8 \pm 49.3	1 (0.4%)
	Crotonaldehyde	HPMMA	623.0 \pm 673.6	300.0 \pm 255.2	1 (0.3%)	1223.9 \pm 789.2	0 (0%)
	N,N-Dimethylformamide	AMCC	299.2 \pm 296.0	160.4 \pm 117.3	5 (1.3%)	527.5 \pm 354.3	0 (0%)

	Ethylbenzene, styrene	PGA	302.9 ± 217.6	225.9 ± 145.5	11 (2.9%)	428.9 ± 254.6	3 (1.3%)
	Propylene oxide	2HPMA	82.9 ± 355.9	58.4 ± 122.9	4 (1.1%)	123.3 ± 555.4	1 (0.4%)
	Styrene, ethylbenzene	PHEMA	2.8 ± 4.1	2.2 ± 4.1	174 (46.4%)	3.9 ± 4.0	37 (16.2%)
		MA	323.0 ± 253.6	232.8 ± 165.5	1 (0.3%)	470.4 ± 299.9	0 (0%)
	Toluene	BMA	11.7 ± 12.4	11.5 ± 12.1	4 (1.1%)	11.9 ± 12.9	0 (0%)
	Xylene	2MHA	54.7 ± 85.5	30.1 ± 53.4	53 (14.1%)	95.0 ± 109.5	6 (2.6%)
		3MHA+4MHA	405.9 ± 539.9	194.9 ± 303.1	30 (8.0%)	750.9 ± 653.4	3 (1.3%)

251 VOC metabolites were measured in 375 non-smoking participants, 25-70 years of age with low
 252 to severe CVD risk. IQR, interquartile range; LOD, limit of detection. Units: µg/g creatinine. The
 253 complete chemical names of each metabolite are listed in Supplemental Table 1.

254

255 VOC metabolites had a high degree of variability in both smokers and non-smokers.

256 Participants who were defined as current smokers consistently showed higher average

257 VOC metabolite levels than non-smokers. Although significant levels of most VOC

258 metabolites could be detected in non-smokers, in many (>20%) samples the level of

259 acrylonitrile (CYMA), styrene/ethylbenzene (PHEMA) metabolites were below the limit of

260 detection (LOD).

261 In single-VOC analysis, the levels of several urinary VOC metabolites were

262 associated with the levels of CACs measured in peripheral blood (Figure 1). In the full

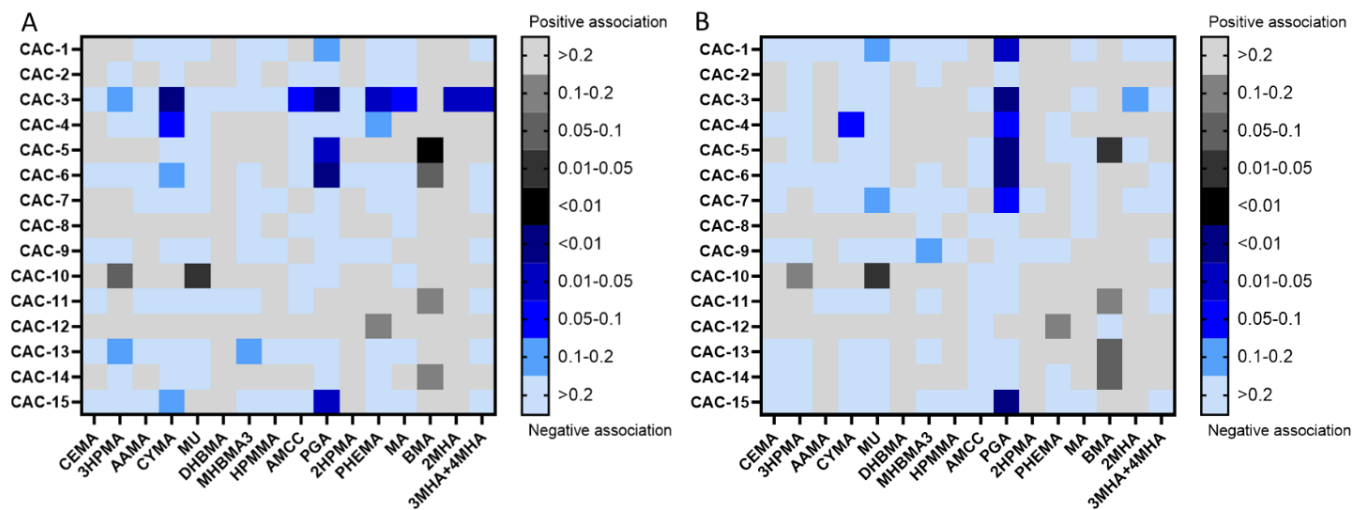
263 data set, PGA was significantly associated with CAC-3, 5, 6, and 15, MU was significantly

264 associated with CAC-10, and BMA was associated with CAC-5. Additionally, we found

265 that CAC-3 was associated with CYMA, PHEMA, 2-MHA, and 3-MHA+4-MHA,

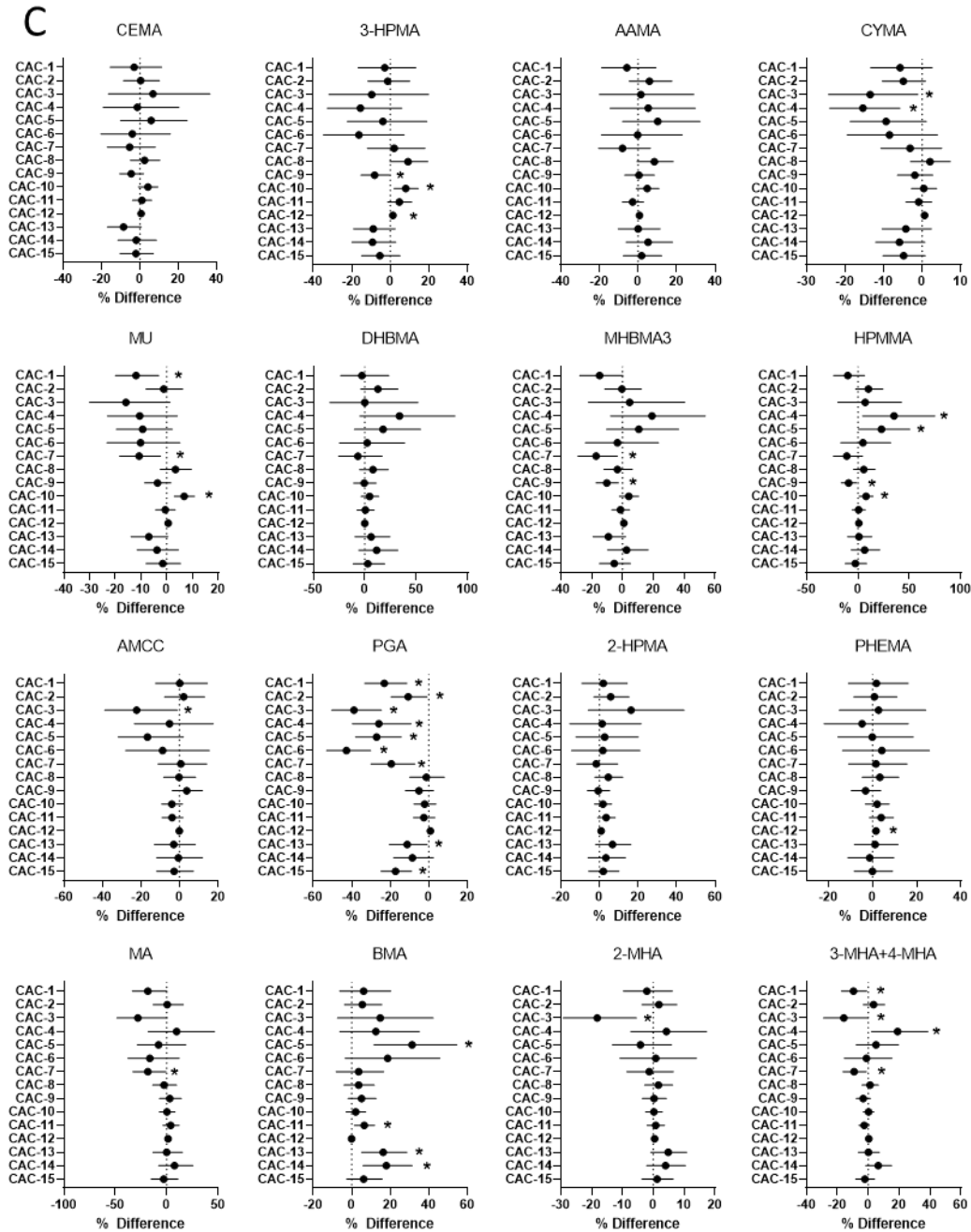
266 suggesting that CAC-3 may be a sensitive marker to tobacco derived VOCs.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



267

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



268

269 **Figure 1.** Association between the urinary metabolites of VOCs and circulating levels of CACs
 270 and heat map of p-values for pair-wise associations between the urinary metabolites of VOCs
 271 and circulating levels of CACs. (A) False Discovery rate adjusted p-values for total sample
 272 (n=603). (B) False Discovery rate adjusted p-values for non-smokers (n=375). Black grids
 273 represent positive associations and blue grids represent negative associations. The intensity of
 274 color represents p-values, with darker grids indicating smaller p-values. (C) The X axis represents
 275 % difference (\pm 95% confidence interval) per 2-fold increase in VOC metabolite for each CAC
 276 sub-population (CAC 1-15) in non-smokers (n=375). Generalized linear models were adjusted for

277 age, sex, race, BMI, diabetes, hypertension, and daily PM_{2.5} levels. Models for the full sample
278 were additionally adjusted for smoking status.

Abbreviations: CEMA, N-Acetyl-S-(2-carboxyethyl)-L-cysteine; 3-HPMA, N-Acetyl-S-(3-hydroxypropyl)-L-cysteine; AAMA, N-Acetyl-S-(2-carbamoylethyl)-L-cysteine; CYMA, N-Acetyl-S-(2-cyanoethyl)-L-cysteine; MU, *trans, trans*-muconic acid; DHBMA, N-Acetyl-S-(3,4-dihydroxybutyl)-L-cysteine; MHBMA3, N-Acetyl-S-(4-hydroxy-2-buten-1-yl)-L-cysteine; HPMA, N-Acetyl-S-(3-hydroxypropyl-1-methyl)-L-cysteine; AMCC, N-Acetyl-S-(N-methylcarbamoyl)-L-cysteine; PGA, Phenylglyoxylic acid ; 2HPMA, N-Acetyl-S-(2-hydroxypropyl)-L-cysteine; PHEMA, N-Acetyl-S-(1-phenyl-2-hydroxyethyl)-L-cysteine +; MA, Mandelic Acid; BMA, N-Acetyl-S-(benzyl)-L-cysteine; 2MHA, 2-Methylhippuric acid; 3MHA+4MHA, 3-Methylhippuric acid + 4-Methylhippuric acid. * Represents significant association.

279

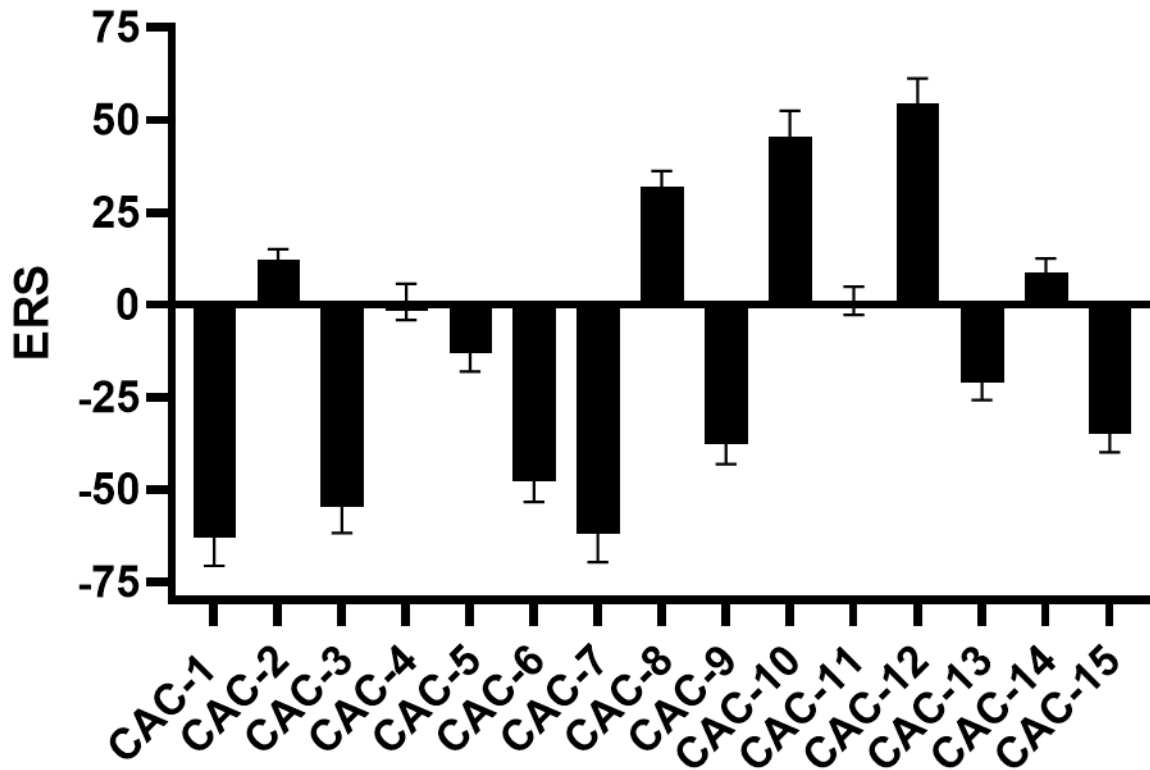
280 **Analysis of non-smokers**

281 In our analysis of non-smokers (Figure 1B, 1C), we found similar associations as
282 in our analysis of the total sample. In non-smokers, some of the most pronounced effects
283 were observed with PGA – a metabolite that could be derived from either ethylbenzene
284 or styrene - which showed a significant inverse association with CACs-1,2,3,4,5,6, and 7,
285 as well as CAC-13 and CAC-15. Several other VOC metabolites showed similar, but
286 variable associations. The percent difference ranged from -10.7% (95% CI: -19.6%, -
287 0.8%) to -42.8% (95% CI: -53.1%, -30.1%). Many VOC metabolites were inversely
288 associated with CAC levels, including metabolites of xylene (3-MHA+4-MHA), acrylonitrile
289 (CYMA), and 1,3-butadiene (MHBMA3). For benzene (MU), we observed a significant
290 inverse association with CAC-1 (-11.9%; 95% CI: -20.0%, -3.0%), and its subpopulation
291 CAC-7 (-10.7%; 95% CI: -18.3%, -2.3%), along with a positive association with CAC-10.
292 Positive associations were also observed with toluene (BMA) and CAC-5, 11, 13, and 14,
293 which share AC133+ marker. After correcting for multiple testing, PGA showed statistically
294 significant associations with CAC-1, 3, 5, 6, and 15 (Figure 1B). Further, MU remained
295 significantly associated with CAC-10 and BMA remained significantly associated with
296 CAC-5.

297 In our sensitivity analysis of non-smokers, additionally adjusting temperature,
298 relative humidity, and ambient ozone did not substantially alter our associations
299 (Supplemental Figure 2). Overall, these observations suggest that exposure to several
300 VOCs is generally associated with lower levels of several CAC subpopulations in
301 peripheral blood, particularly the early progenitor CACs.

302 Because we found similar associations between CACs and VOCs in the total
303 sample, and in non-smokers, we focused the rest of our analysis on the subgroup of non-
304 smokers, to determine how low levels of VOCs are associated with endothelial damage.
305 In non-smokers, there was minimal correlation between VOC metabolites and average
306 daily ambient PM_{2.5}, ozone, temperature, and relative humidity (Supplemental Figure 1).
307 The levels of urinary VOC metabolites showed a strong dependence on demographics.
308 A comparison of urinary metabolites showed that females had higher levels of the
309 metabolites of N,N'-DMF (AMCC), ethylbenzene/styrene (phenylglyoxylic acid, PGA and
310 mandelic acid, MA) and toluene (BMA) than males. Likewise, in comparison with Blacks,
311 Whites had higher levels of metabolites of benzene (*trans, trans*-muconic acid, MU), N,N'-
312 DMF (N-Acetyl-S-(N-methylcarbamoyl)-L-cysteine, AMCC) and styrene/ethylbenzene (MA).
313 The metabolites of several VOCs – 1,3, butadiene (N-Acetyl-S-(3,4-dihydroxybutyl)-L-
314 cysteine, DHBMA), crotonaldehyde (N-Acetyl-S-(3-hydroxypropyl-1-methyl)-L-cysteine,
315 HPMMA), N,N'-DMF (AMCC), styrene/ethylbenzene (MA), toluene (BMA), and xylene
316 (3MHA+4MHA) were also higher in those >55 years of age. Urinary levels of the
317 metabolite of butadiene (MHBMA3) and toluene (N-Acetyl-S-(benzyl)-L-cysteine, BMA and,
318 2-Methylhippuric acid, 2MHA) levels were higher in those with a BMI of <30 (Supplemental
319 Table 2), reflecting different levels of exposure.

320 To determine the extent to which total VOC exposure can affect CAC levels, and
321 which cell populations contribute to the directionality of the associations (found in Figure
322 1) we created an Environmental Risk Score (ERS) for each cell population (Figure 2).



323

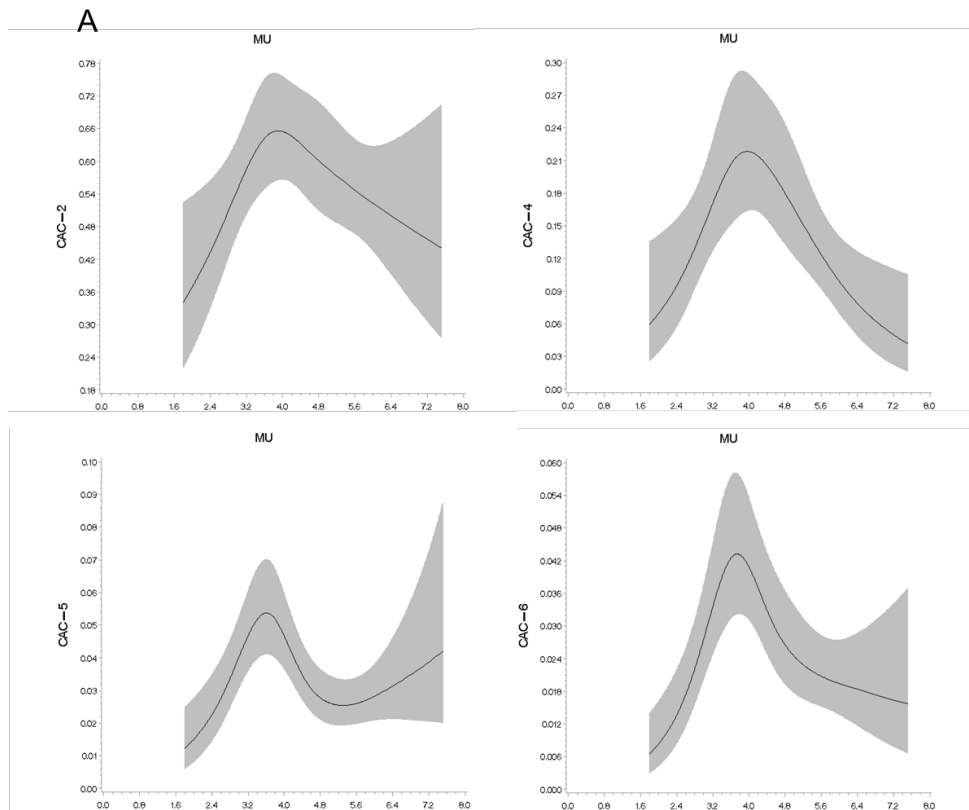
324 **Figure 2.** Association between Environmental Risk Scores and Circulating Angiogenic Cells in
325 non-smokers. Environmental Risk score was calculated for each cell population as follows: 1.
326 The urinary levels of VOC metabolites were log transformed. 2. Each log transformed VOC
327 metabolite level was regressed against the levels of the cells separately, while adjusting for
328 covariates. 3. ERS was calculated as the sum of the log transformed VOC metabolite levels x
329 (regression coefficient/standard error). Mean and standard deviation of the ERS was calculated
330 for each CAC. Values are mean \pm SD of environmental risk score for each cell population.

331

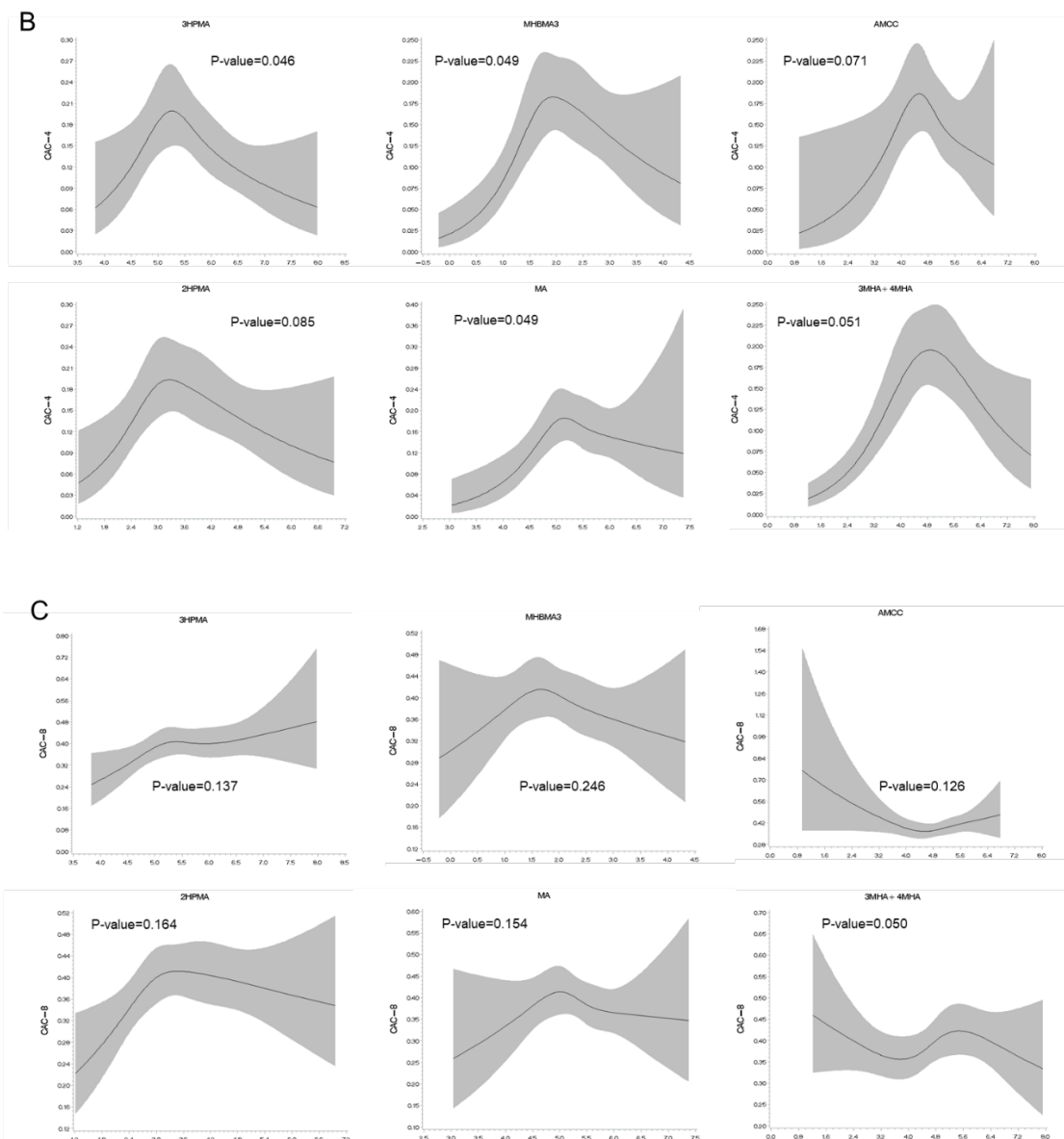
332 We found a strong negative risk score for CAC-1 (-62.9), CAC-3 (-54.6), and CAC-
333 7 (-62.1), which are a related group of cells with common markers
334 CD45^{dim}/CD146⁺/CD34⁺. To a lesser extent, we also found negative risk scores for CAC-
335 6 (-47.6), and CAC-9 (-37.7), which are CD34⁺ cells. In comparison, we found positive

336 risk scores for CD45⁺ cells (CAC-12), along with subgroups of CD45⁺ cells: CAC-2, CAC-
337 8, and CAC-14. Moreover, we also found a positive score for CAC-10, which is CD146⁺.
338 This indicates that the levels of CD45⁺ cells were positively associated with the urinary
339 metabolites of VOC metabolites, whereas CD45^{dim} and CD34⁺ cells were negatively
340 associated with the metabolite levels.

341 We used restricted cubic splines to determine whether the relationship between
342 VOC metabolites and CACs was non-linear (Figure 3).



343



344

345

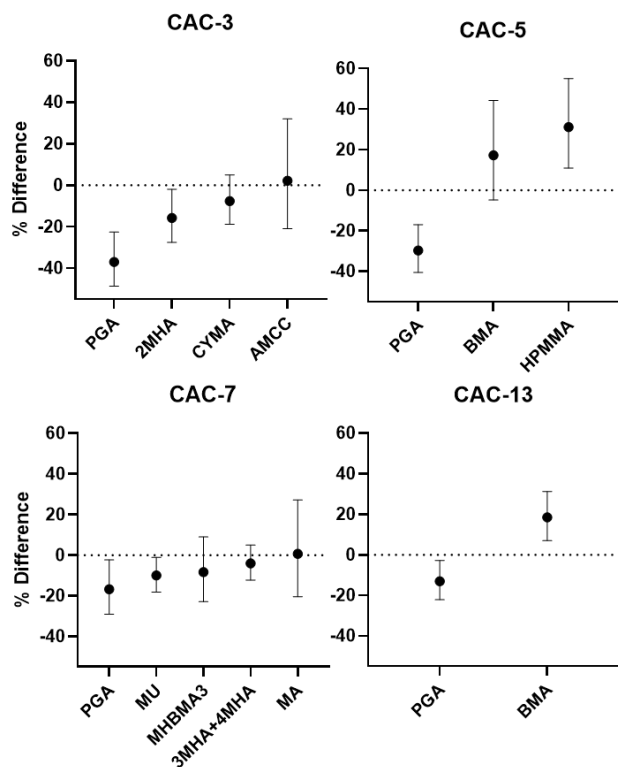
346 **Figure 3.** Exposure-response relationships between selected CACs and VOCs in non-smokers.
347 Non-linear relationships between MU metabolites and CACs (A); VOCs and CAC-4 (B); and
348 VOCs and CAC-8 (C) were examined using restricted cubic splines. Tests for non-linearity used
349 the likelihood ratio test. Models are adjusted for age, sex, race, BMI, diabetes, hypertension, beta-
350 blocker use, and daily PM_{2.5} levels. P-value represents likelihood ratio test for non-linearity.
351

352 We observed non-linear relationships between the circulating levels of several cell
353 populations and the metabolite of benzene (MU). In general, low levels of MU was
354 associated with higher levels of circulating cells, whereas higher MU levels were

355 associated with lower cell levels. These non-linear relationships are presented in Figure
356 3A, which shows a near significant ($p < 0.1$) non-linear relationships with CD146+ cell
357 populations (CAC-2, 4, 5, and 6), and with AC133+ cell population – CAC-4 and CAC-5.

358 To determine whether AC133+ cells contribute to the non-linear relationships, we
359 examined the association of subgroups of CD146+ cells, CAC-4 (CD45+/CD146+/AC133+)
360 and CAC-8 (CD45+/CD146+/AC133-) with VOC metabolites. We found near significant
361 ($p < 0.1$) non-linear relationships between CAC-4 and metabolites of acrolein, 1,3-
362 butadiene, N, N-dimethylformamide, propylene oxide, ethylbenzene/styrene, and xylene.
363 These non-linear relationships were not observed with CAC-8, with the exception of
364 xylene (3-MHA+4-MHA), suggesting that early progenitor endothelial cells
365 (CD146+/AC133+) are recruited at low exposure levels, but are depleted at high VOC
366 exposures.

367 Given that co-exposure to VOCs is common, we used multi-VOC models to assess
368 whether the significant VOC effects found in single-VOC models were confounded by co-
369 exposures to other VOCs (Figure 4).



370

371 **Figure 4.** Association between circulating angiogenic cells and multiple VOCs in the same model
372 (non-smokers). Values represent % difference per 2-fold increase in VOC metabolite. VOC
373 metabolites are ordered by strength of their regression coefficients. Only significant metabolites
374 from one-pollutant models were entered in the model. Models adjusted for age, sex, race, BMI,
375 diabetes, hypertension, and daily PM_{2.5} levels.
376

377 For this analysis, we focused on rare cell populations with negative ERS scores CAC-3,
378 CAC-5, CAC-7, and CAC-13 (Figure 2). We observed that the metabolite of
379 ethylbenzene/styrene (PGA) remained negatively associated with CAC-3 (-37.0% per 2-
380 fold increase; 95% CI: -48.7%, -22.6%), CAC-5 (-29.8%; 95% CI: -40.6%, -17.0%), CAC-
381 7 (-16.8%; 95% CI: -29.1%, -2.4%), and CAC-13 (-13.1%; 95% CI: -22.2%, -2.8%),
382 suggesting that the associations between PGA and CACs are independent of other VOC
383 metabolite levels. We also found that the metabolites of xylene (2MHA) remained
384 negatively associated with CAC-3, crotonaldehyde (HPMMA) remained positively

385 associated with CAC-5, benzene (MU) remained negatively associated with CAC-7, and
386 toluene (BMA) remained positively associated with CAC-13.

387 In our Bayesian kernel machine regression analysis, important VOC metabolites
388 contributing to changes in CACs are indicated by high posterior inclusion probabilities
389 (Table 3).

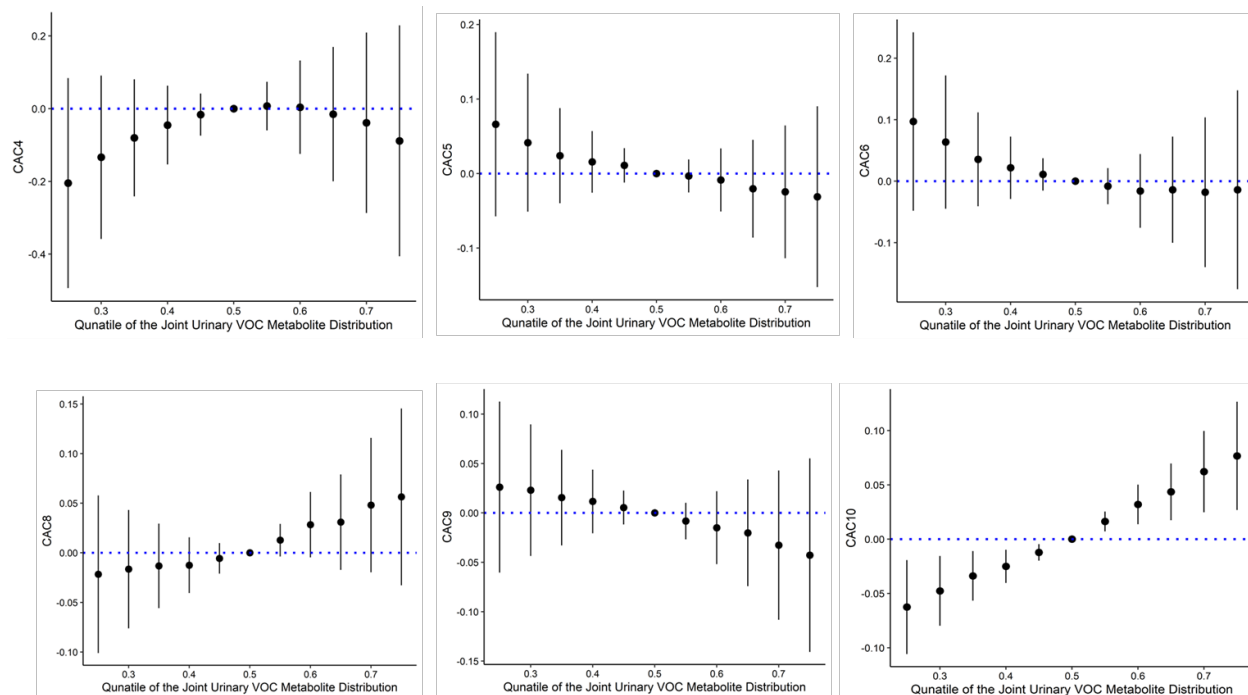
390 **Table 3.** Posterior inclusion probabilities of individual VOC metabolites in the BKMR
391 model for 15 CACs.

VOC	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CEMA	0.53	0.09	0.64	0.74	0.00	0.03	0.49	0.03	0.36	0.00	0.00	0.00	0.11	0.42	0.41
3HPMA	0.43	0.26	0.62	0.68	0.00	0.02	0.54	0.16	0.50	0.13	0.30	0.00	0.15	0.45	0.35
AAMA	0.43	0.10	0.61	0.65	0.00	0.00	0.48	0.06	0.35	0.00	0.00	0.00	0.04	0.42	0.39
CYMA	0.47	0.05	0.68	0.68	0.00	0.01	0.52	0.05	0.38	0.00	0.07	0.00	0.08	0.42	0.34
MU	0.48	0.13	0.60	0.77	0.00	0.01	0.44	0.06	0.38	0.55	0.00	0.00	0.18	0.38	0.39
DHBMA	0.57	0.12	0.61	0.78	0.00	0.02	0.58	0.01	0.33	0.00	0.02	0.00	0.02	0.45	0.47
MHBMA3	0.47	0.10	0.61	0.78	0.01	0.00	0.50	0.01	0.51	0.00	0.00	0.00	0.33	0.35	0.33
HPMMA	0.46	0.14	0.59	0.74	0.01	0.06	0.50	0.01	0.46	0.03	0.01	0.00	0.03	0.36	0.35
AMCC	0.50	0.10	0.57	0.70	0.03	0.01	0.58	0.03	0.35	0.00	0.03	0.00	0.08	0.46	0.39
PGA	0.54	0.19	0.64	0.62	0.03	0.02	0.80	0.03	0.26	0.00	0.02	0.00	0.04	0.44	0.44
2HPMA	0.43	0.17	0.61	0.70	0.00	0.01	0.53	0.04	0.30	0.02	0.05	0.00	0.04	0.37	0.35
PHEMA	0.47	0.07	0.65	0.61	0.01	0.15	0.49	0.05	0.40	0.03	0.00	0.00	0.06	0.40	0.34
MA	0.55	0.12	0.72	0.73	0.02	0.02	0.53	0.04	0.45	0.00	0.00	0.00	0.07	0.49	0.35
BMA	0.42	0.10	0.61	0.71	0.00	0.00	0.46	0.02	0.37	0.01	0.02	0.00	0.91	0.57	0.35
2MHA	0.53	0.19	0.76	0.76	0.02	0.00	0.74	0.02	0.38	0.00	0.06	0.00	0.09	0.34	0.41
3MHA+4MHA	0.46	0.10	0.62	0.89	0.12	0.03	0.71	0.01	0.30	0.00	0.01	0.00	0.05	0.44	0.37

392 Models (n=375) were adjusted for age, sex, race, BMI, diabetes, hypertension, and daily PM_{2.5} levels.

393 The directionalities of single VOC associations in our mixture models were generally
394 consistent with the observed directionalities in our single VOC generalized linear models.
395 For example, in mixture models, individuals with greater PGA levels had lower numbers
396 of CACs, while individuals with greater BMA and HPMMA levels had higher numbers of
397 CACs. As evident in Figure 5, we also found that the mixture of VOCs had a nonlinear
398 relationship with CAC-4, similar to the nonlinear associations observed in our single VOC

399 models. We also found that greater exposure to the VOC mixture was associated with
400 higher CAC-10 levels (Figure 5).

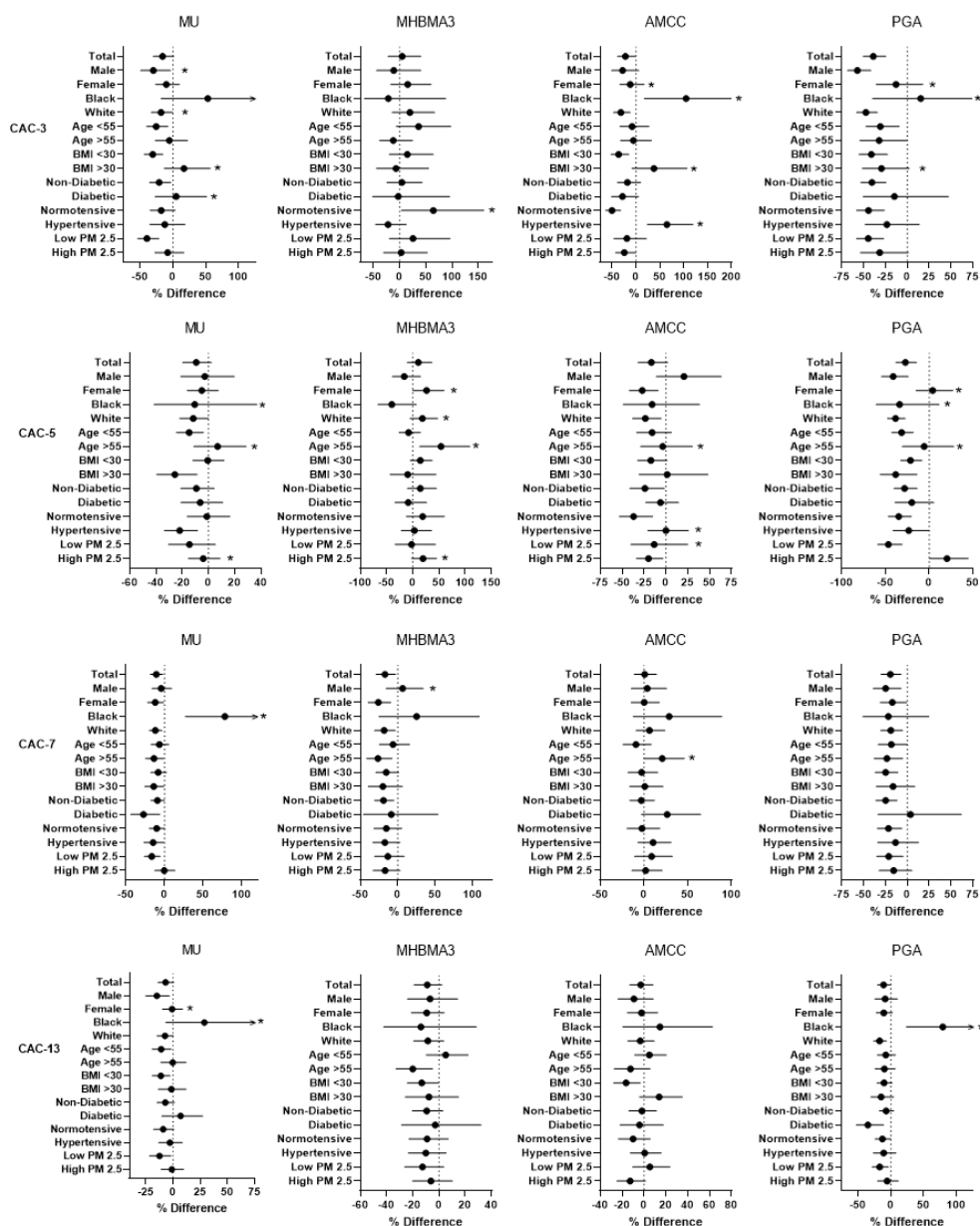


401
402 **Figure 5.** Overall association of the VOC metabolite mixture (and 95% credible intervals) with
403 selected CACs, when all exposures are set to a given percentile relative to median. Models (non-
404 smokers; n=375) were adjusted for age, sex, race, BMI, diabetes, hypertension, and daily PM_{2.5}
405 levels.
406

407 The effect of the VOC mixture on CAC-10 was largely driven by benzene (MU), with a
408 posterior inclusion probability of 0.55, where all other VOC metabolites were less than
409 0.15. We found minimal evidence of interactions between pairs of VOC metabolites in our
410 mixture analysis (data not shown).

411 Because the metabolites of benzene, 1,3-butadiene, N,N-dimethylformamide, and
412 ethylbenzene/styrene consistently showed strong associations with CACs, we
413 investigated whether these associations were modified by the demographic
414 characteristics of the participants or their co-morbidity profile. In our subgroup analysis,

415 we found a common theme in many of the models, in which participants who were male,
 416 White, normotensive, and recruited on low PM_{2.5} days showed stronger inverse
 417 relationships between VOCs and CAC levels (Figure 6).



418
 419 **Figure 6.** Subgroup analysis of the relationship between different CAC population with the urinary
 420 metabolites of benzene (MU), 1,3-butadiene (MHBMA3), ethylbenzene/styrene (PGA), and N,N-
 421 Dimethylformamide (AMCC) in non-smokers. Generalized linear models were adjusted for

422 covariates - age, sex, race, BMI, diabetes, hypertension, and daily PM_{2.5} levels, unless the
423 covariates were used for stratification. *Significant interaction between subgroups in the full
424 model.

425

426 For example, for CAC-13, we found a negative association between
427 ethylbenzene/styrene metabolite (PGA) in White participants of -17.2% per 2-fold
428 increase in CAC levels (95% CI: -26.4%, -7.0%), whereas in Black participants, we found
429 a positive association of 79.1% (95% CI: 23.7%, 159.3%). We observed a similar
430 relationship for the metabolite of benzene (MU), when stratified by race; however, this
431 may be due to lower levels of MU in urine samples from Black participants (62.2 µg/g
432 creatinine) than White participants (146.0 µg/g creatinine), as presented in *Supplemental*
433 *Table 2*.

434 We also found some evidence that age and BMI modify the relationships between
435 VOCs and CAC levels, although the direction was not consistent in all models.
436 Interestingly, in several of our models, we found stronger inverse associations between
437 VOC metabolite and CACs in normotensives than hypertensives, and to a lesser extent,
438 non-diabetics than diabetics, - conditions that are known to affect cell migration. For
439 example, for CAC-5, we found a negative association with ethylbenzene/styrene
440 metabolite (PGA) in non-diabetics of -28.2% per 2-fold increase, and with N,N-
441 dimethylformamide (AMCC) in non-diabetics of -24.0% per 2-fold increase, while the
442 relationships in diabetics were not significant. To determine whether these results were
443 due to differential levels of VOC exposure between subgroups, we calculated average
444 metabolite levels for each subgroup *Supplemental Table 2*. We found that levels of MU
445 (benzene) and PGA (ethylbenzene and styrene) were higher in diabetics than non-

446 diabetes and were higher in hypertensive than normotensive individuals, although
447 normotensive had higher levels of 2HPMA (propylene oxide).

448 **4. DISCUSSION**

449 The major findings of this study are that the urinary metabolites of several common
450 VOCs, such as ethylbenzene/styrene, benzene, xylene, and 1,3-butadiene were
451 negatively associated with circulating angiogenic cells. We found similar associations in
452 non-smokers compared with the total sample containing smokers. For some VOCs, such
453 as benzene, we observed that low levels of their metabolites were positively associated
454 with an increase in CAC levels, whereas at high levels of exposure were associated with
455 a significant decrease in CAC levels. Our environmental risk score identified that CAC-
456 1, CAC-3, and CAC-7, with common cell markers CD45^{dim}/CD146⁺/CD34⁺, were the most
457 sensitive to total VOC exposure. Our mixture analysis revealed that the total VOC
458 metabolite mixture was positively associated with CD146⁺ cells. We also found sex, race,
459 age, and hypertension to be effect modifiers of the some of the relationships between the
460 levels of VOC metabolites and CACs. Taken together, these findings suggest that
461 exposure to VOCs could lead to deficits in endothelial repair capacity in a dose-dependent
462 manner, which vary with sex, race, and comorbidities. Given that low CAC levels are
463 predictive of future cardiovascular events and death ³⁹, these results support the notion
464 that VOC exposure is associated with significant vascular injury independent of tobacco
465 use, and therefore it may be an important environmental risk factor for CVD.

466 The potential cardiovascular effects of VOCs derived from tobacco smoke have
467 been extensively studied. ⁹⁻¹¹ We found similar associations between VOCs and CACs in
468 our total sample containing smokers, compared with our subgroup of non-smokers. This

469 suggests that the effects of VOC exposure on CACs is independent of tobaccos smoke
470 exposure and is still present at environmentally low levels. We additionally found that
471 several VOC metabolites were associated with CAC-3 in the total sample analysis. We
472 have previously shown that low levels of CAC-3 is associated with higher prevalence of
473 diabetes, higher HbA1c levels, and impaired vascular function.³⁷ This could be due to the
474 higher levels of several VOC metabolites in the total sample, such as CYMA
475 (acrylonitrile), and PHEMA (ethylbenzene/styrene), which are typically difficult to detect
476 in non-smokers. In contrast, the cardiovascular effects of VOCs not derived from tobacco
477 smoke have received scant attention. Nonetheless, in addition to tobacco smoke,
478 automobile emissions, industrial facilities, hazardous waste sites, and household
479 products are major sources of VOC emissions, and therefore, VOC exposures are
480 common and frequent.¹³ In our study, we found significant levels of the VOC metabolites
481 in the urine of non-smokers, indicating particularly high levels of exposure to acrolein,
482 crotonaldehyde, and 1,3-butadiene, and acrylamide, as well as the BTEXS (benzene,
483 toluene, ethylbenzene, xylene and styrene) group of VOCs, suggesting a contribution of
484 both indoor and outdoor exposure. Specifically, we found that the CAC levels were
485 particularly sensitive to metabolites of BTEXS compounds. We observed significant
486 associations between CACs and urinary metabolite of benzene (3 CACs), toluene (4
487 CACs), ethylbenzene/styrene (9 CACs), and xylene (4 CACs).

488 Although exposure to BTEXS is common, little is known about their cardiovascular
489 effects. A study from Hong Kong found that short term elevations in atmospheric levels
490 of benzene and TEX (toluene, ethylbenzene, and xylene) was associated with increases
491 in circulatory disease mortality by 5.8% and 3.5%,⁵¹ although in a study from Toronto, no

492 intra-urban variations in VOC levels were found be associated with cardiovascular
493 disease mortality. ¹⁶ Other BTEXS such as toluene, styrene, ethylbenzene, and xylenes
494 have been also linked to a higher prevalence of CVD, ⁵² and exposure to VOCs such as
495 benzene, ethylbenzene, and styrene have also been found to be associated with higher
496 odds of developing metabolic syndrome. ⁵³ Along with these data, the results of our study,
497 support the notion that exposure to VOC could have adverse cardiovascular
498 consequences, potentially due to low-level vascular injury reflected by changes in CAC
499 levels.

500 We found inverse associations between metabolites of acrylonitrile and 1, 3-
501 butadiene and CACs. Although the cardiovascular effects of acrylonitrile are less well
502 studied, exposure to 1,3-butadiene has been found to accelerate the formation of
503 atherosclerotic lesions in animal models, and polymorphisms in phase II enzymes that
504 metabolize 1,3-butadiene are associated with the development of atherosclerosis in
505 humans. ⁵⁴ Hence, our observation that urinary metabolites of 1,3-butadiene are
506 associated with CAC levels suggest that the cardiovascular risk of exposure to this VOC
507 may be in part due to depletion of early, progenitor cells, which aid in vascular repair and
508 thereby can prevent the progression of atherosclerotic lesions. Therefore, our
509 observations linking exposure to VOCs such as 1,3-butadiene to concurrent changes in
510 CACs, not only provides a new biological basis for the cardiovascular effects of VOCs,
511 independent of concurrent exposure to PM_{2.5}, but also provides new insights into their
512 mechanism of toxicity.

513 In real world scenario, the urban air contains many different VOCs and therefore,
514 we conducted multi-VOC models controlling for all significantly associated metabolites

515 from the single VOC models. In particular, we found that the urinary metabolite PGA
516 remained significantly associated with the levels of CD45^{dim}/CD146⁺/CD34⁺ cells (CAC-
517 3, CAC-7), AC133⁺/CD146⁺ cells (CAC-5), and AC133⁺/CD34⁺ cells (CAC-13). PGA is
518 oxidized from mandelic acid (MA) by alcohol dehydrogenase, and both are considered
519 the major urinary metabolites of styrene and ethylbenzene exposure.⁵⁵ Styrene is mainly
520 used in the manufacturing of plastics, synthetic rubbers, and resins, and is the top
521 released carcinogen in the United States,⁵⁶ while ethylbenzene is found naturally in coal
522 tar and petroleum, and used in the manufacturing of styrene.⁵⁵ The primary route of
523 exposure to styrene and ethylbenzene is through inhalation, contributing up to 99% of
524 total exposure, with 1-2% from food consumption.⁵⁷ We found that the
525 ethylbenzene/styrene metabolite - PGA, but not MA was negatively associated with
526 CACs. This could be due to the chemical structure of PGA, which is a ketone, and is
527 therefore, more reactive than the acid metabolite MA. This could also be representative
528 of higher ethylbenzene/styrene exposure, which would affect the ratio of PGA to MA
529 through an increase in alcohol dehydrogenase activity. The results of our multi-VOC
530 analysis indicate that CACs may be particularly sensitive to ethylbenzene/styrene
531 exposures.

532 We found that the negative association between the N,N-DMF metabolite (AMCC)
533 and CACs was significantly stronger in normotensive participants than those with
534 hypertension. Previous work has shown that hypertension is associated with lower CAC
535 levels and that hypertension is a major independent predictor of impaired endothelial
536 progenitor cell migration.⁵⁸ It has also been demonstrated that hypertension accelerates
537 EPC senescence in both clinical and experimental settings.⁵⁹ These findings suggest

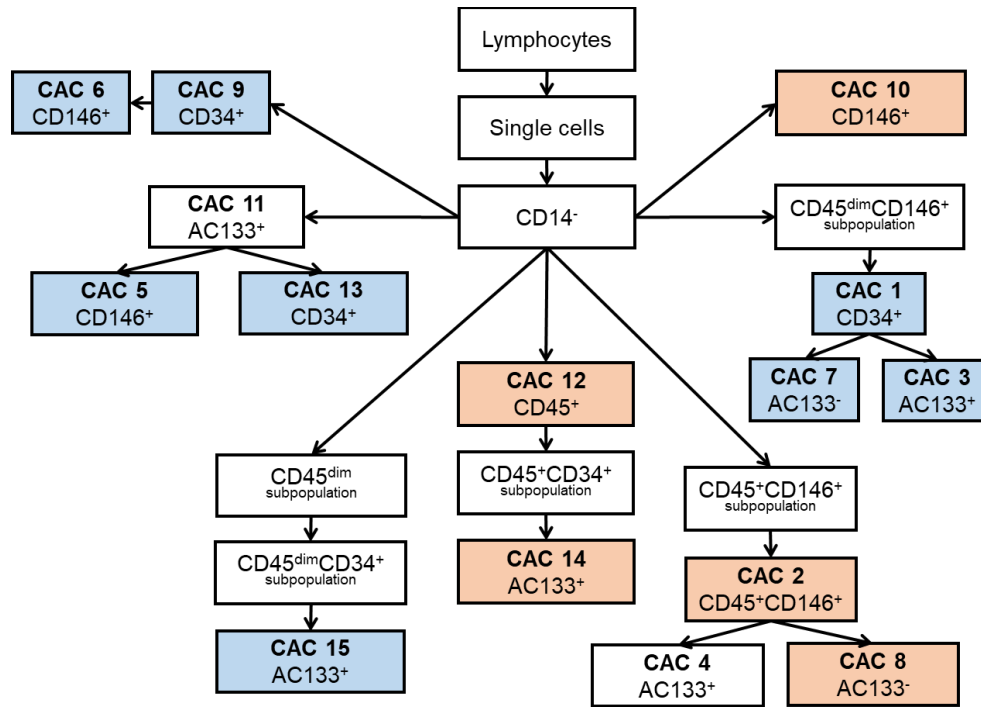
538 that because of the inherent susceptibility of CACs to hypertension, CACs in hypertensive
539 participants may be more vulnerable to exposure to VOCs such as N,N-DMF. N,N-DMF
540 is a commonly used solvent with wide industrial applications, and occupational exposure
541 to high levels of N,N-DMF is associated with liver injury;⁶⁰ however, our observation that
542 even ambient exposure to N,N-DMF is associated with vascular injury, suggests that
543 there may be a high risk of hypertension and cardiovascular disease even with low-level
544 exposures.

545 The observed association between VOC metabolites and diminished circulating
546 angiogenic cells suggest that individuals exposed to VOCs may have a lower capacity to
547 repair the endothelium. These effects of VOC exposure appear to be similar to those of
548 other environmental pollutants, which are also associated with suppression of CAC levels.
549 Previous work has shown that exposure to PM_{2.5} over a period of 24h to several weeks
550 suppresses EPC levels,^{41, 61} and exposure to pollutants such as particulate nickel,⁶²
551 acrolein¹⁷ and benzene¹⁸ have also been shown to reduce CAC levels. These
552 observations suggest that pollutant exposure either depletes CAC levels in peripheral
553 blood or interferes with the release from the bone marrow and the spleen. However, in
554 contrast to such suppression, CAC levels are slightly elevated in individuals breathing
555 coarse PM⁶³ or those living close to a major roadway.⁶⁴ This biphasic response is
556 consistent with the dose-response relationships observed in the current study, which are
557 indicative of increased CAC mobilization (to promote repair) at low levels of exposure,
558 and CAC suppression at higher exposures.

559 Recently, we showed that living in areas of high vegetation is associated with
560 increased levels of CD34⁺/CD45⁺/AC133⁺ CACs, but decreased

561 CD31⁺/CD34⁺/CD45^{dim}/AC133⁺ CACs, ⁶⁵ suggesting greenness may be protective
562 against air pollution related health effects. Thus, it seems that low levels of toxicity could
563 lead to increased recruitment of CAC into the peripheral blood, whereas high levels of
564 exposure could suppress CAC recruitment and/or depleting CAC by increasing utilization.
565 The directionality of the effects of exposure on CAC levels may also depend upon locus
566 of injury, some exposures may prevent recruitment, while others may induce injury that
567 stimulates recruitment. Regardless, our results with VOCs suggest that exposure to
568 these chemicals suppresses CAC levels in peripheral blood, but further studies are
569 required to determine whether this is due to decreased recruitment or increased
570 utilization.

571 We found that several VOC metabolite levels were negatively associated with
572 angiogenic cells (CD45^{dim}/CD146⁺/CD34⁺), but positively associated with hematopoietic
573 endothelial cells (CD45⁺/CD146⁺), as evident from our environmental risk score results
574 (Schematic 2).



575

576 **Scheme 2.** Gating Strategy for the Human Circulating Angiogenic Cells. Blue indicates negative
577 environmental risk score; orange indicates positive environmental risk score.

578

579 In comparison with CD45^{dim} cells, CD45⁺ CACs are more primitive, with the potential to
580 self-renew and differentiate into specialized cells that can control biological activities such
581 as homeostasis and response to inflammation.⁶⁶ Additionally, the angiogenic cells are
582 rare, and can be easily depleted when used in the repair of endothelial injury. Our
583 exposure response curves are consistent with this notion that at low levels of VOC
584 exposure, CD45⁺ cells have the ability to regenerate to repair endothelial injury, however
585 at high VOC exposure, the injury surpasses the cells capacity to regenerate or the
586 mobilization of these cells from depots such as the spleen or bone marrow is prevented.
587 Consistent with this notion, results from our Bayesian kernel machine regression also
588 found a nonlinear relationship with the mixture of VOCs and CAC-4
589 (CD45⁺/CD146⁺/AC133⁺). Further, we found that the mixture of VOC metabolites_{was}

590 positively associated with CD146⁺ cells. These cells are thought to be mature cells which
591 detach from the intimal monolayer after endothelial injury. Studies have shown an
592 increased number of CD146⁺ cells in CVD, such as unstable angina, acute myocardial
593 infarction, stroke, and diabetes.⁶⁷ Positive associations could also be due to the specific
594 VOC exposure. We found primarily positive relationships between the toluene metabolite
595 (BMA) and CACs. This could be reflective of the low level of toxicity of toluene in
596 comparison to other VOCs.

597 The major strength of this study is the use of individual level data to assess
598 exposure. Although, measurements of the urinary metabolites do not permit source
599 assignment, they do provide an integrated measure to total exposure of an individual
600 whether via inhalation, ingestion, or endogenous production. Hence, in linking exposure
601 to outcome, such measurements provide a unique advantage as exposure-response
602 could be concurrently evaluated in the same individual. Using this approach, we were
603 able to assess association between exposure and a sensitive and robust measure of
604 preclinical cardiovascular injury (CACs) which is predictive of CVD events and mortality.
605 In our analysis, we also adjusted for exposure to PM_{2.5} and were able to apply a
606 composite VOC risk score that could be used in future studies.

607 Our study also has some limitations. Because of the cross-sectional nature of the
608 study, the associations found in this study cannot address causality. Longitudinal studies
609 are needed to determine whether VOC exposures are associated with adverse clinical
610 outcomes, and whether these outcomes are related to changes in CAC levels. Our study
611 also has the potential for type I errors due to testing of multiple CACs and VOC
612 metabolites. However, due to the correlation between VOC metabolites, and limited prior

613 research on VOC exposure and CVD, we did not adjust for multiple comparisons, to
614 preserve our type II error, and identify potential VOC candidates for future studies.
615 Moreover, we used spot urine collection from each individual to estimate VOC exposure.
616 Although, it has been reported that the inter-day reproducibility for most urinary VOC
617 metabolites is good to excellent.^{68, 69} Following exposure, many VOCs rapidly undergo
618 biotransformation to generate metabolites that can conjugate with glutathione. Other
619 enzymatic reactions remove glutamic acid and glycine to produce a cysteine conjugate,
620 which is N-acetylated and excreted in the urine as mercapturic acid.¹³ It has also been
621 shown that smoking one cigarette, which generates high levels of VOCs, is sufficient to
622 immediately affect the levels of CACs.⁷⁰ Therefore, because VOC metabolites and CAC
623 levels were concurrently measured, it is likely that the observed association represent
624 consequences of recent exposure. Additionally, our urinary VOC metabolites were
625 normalized to creatinine to account for dilution differences, which could introduce error
626 that relates to the dependence of creatinine excretions due to muscle mass, physical
627 activity, and kidney disease. Another limitation of our study is that most participants were,
628 by study design, of older age and at increased risk of cardiovascular disease. Thus, our
629 sample could represent a group of susceptible individuals, which may limit generalizability
630 of our findings.

631 In conclusion, we showed that metabolites of several VOCs were negatively
632 associated with circulating levels of CD45^{dim}/CD146⁺/CD34⁺ cells and the mixture of VOC
633 metabolites was positively associated with CD146⁺ cells in the peripheral blood. These
634 findings suggest that exposure to low ambient levels of VOCs could damage the
635 endothelium, and deplete rare CACs or inhibit their recruitment, and thereby compromise

636 endothelial repair. Given the global burden of CVD, and the ubiquitous distribution of
637 VOCs, it is imperative to further assess their contribution to CVD risk. Because several
638 commonly used products, such as paints, self-care products, and glue, are important
639 source of VOCs, and that there are several sources of indoor emissions (compressed
640 wood, fire-retardants), this study raises the possibility that environmental exposure to
641 VOCs may be an important determinant of CVD risk. Our work also underscores the need
642 for greater awareness of the potential dangers of VOC exposure and heightened vigilance
643 against such exposures. Future longitudinal and mechanistic studies are needed to
644 confirm current findings. If substantiated, an association between VOC exposure and
645 CVD risk could be included in estimates of CVD risk and appropriate limits and
646 recommendations could be developed to reduce both ambient and indoor VOC
647 exposures.

648 **Source of Funding:** This research was supported in part by NIH grants ES023716,
649 ES029846, ES019217, HL149351, HL120163 and GM127607; and Jewish Heritage
650 Foundation of Excellence grant GN190574L.

651 **Acknowledgements:** The authors would like to thank all the partners, members, and
652 volunteers involved in the HEAL study.

653

654 *The authors declare they have no actual or potential competing financial interests.*

655

656

References

- 657 1. Cohen AJ, Brauer M, Burnett R, Anderson HR, Frostad J, Estep K, Balakrishnan K,
658 Brunekreef B, Dandona L, Dandona R, Feigin V, Freedman G, Hubbell B, Jobling A, Kan
659 H, Knibbs L, Liu Y, Martin R, Morawska L, Pope CA, Shin H, Straif K, Shaddick G, Thomas
660 M, van Dingenen R, van Donkelaar A, Vos T, Murray CJL, Forouzanfar MH. Estimates
661 and 25-year trends of the global burden of disease attributable to ambient air pollution: an
662 analysis of data from the Global Burden of Diseases Study 2015. *The Lancet*
663 2017;**389**:1907-1918.
- 664 2. Yang B-Y, Qian Z, Howard SW, Vaughn MG, Fan S-J, Liu K-K, Dong G-H. Global
665 association between ambient air pollution and blood pressure: A systematic review and
666 meta-analysis. *Environmental Pollution* 2018;**235**:576-588.
- 667 3. Mustafić H, Jabre P, Caussin C, et al. Main air pollutants and myocardial infarction: A
668 systematic review and meta-analysis. *Jama* 2012;**307**:713-721.
- 669 4. Brook RD, Cakmak S, Turner MC, Brook JR, Crouse DL, Peters PA, Van Donkelaar A,
670 Villeneuve PJ, Brion O, Jerrett M. Long-term fine particulate matter exposure and mortality
671 from diabetes mellitus in Canada. *Diabetes care* 2013;DC_122189.
- 672 5. Rajagopalan S, Brook RD. Air pollution and type 2 diabetes: mechanistic insights.
673 *Diabetes* 2012;**61**:3037-3045.
- 674 6. Shah ASV, Lee KK, McAllister DA, Hunter A, Nair H, Whiteley W, Langrish JP, Newby DE,
675 Mills NL. Short term exposure to air pollution and stroke: systematic review and meta-
676 analysis. *BMJ* 2015;**350**.
- 677 7. Rajagopalan S, Al-Kindi SG, Brook RD. Air Pollution and Cardiovascular Disease: JACC
678 State-of-the-Art Review. *Journal of the American College of Cardiology* 2018;**72**:2054-
679 2070.
- 680 8. Brook RD, Rajagopalan S, Pope III CA, Brook JR, Bhatnagar A, Diez-Roux AV, Holguin
681 F, Hong Y, Luepker RV, Mittleman MA. Particulate matter air pollution and cardiovascular
682 disease: an update to the scientific statement from the American Heart Association.
683 *Circulation* 2010;**121**:2331-2378.
- 684 9. Lorkiewicz P, Riggs DW, Keith RJ, Conklin DJ, Xie Z, Sutaria S, Lynch B, Srivastava S,
685 Bhatnagar A. Comparison of Urinary Biomarkers of Exposure in Humans Using Electronic
686 Cigarettes, Combustible Cigarettes, and Smokeless Tobacco. *Nicotine & tobacco*
687 *research : official journal of the Society for Research on Nicotine and Tobacco*
688 2019;**21**:1228-1238.
- 689 10. De Jesús VR, Bhandari D, Zhang L, Reese C, Capella K, Tevis D, Zhu W, Del Valle-Pinero
690 AY, Lagaud G, Chang JT, van Bemmelen D, Kimmel HL, Sharma E, Goniewicz ML, Hyland
691 A, Blount BC. Urinary Biomarkers of Exposure to Volatile Organic Compounds from the
692 Population Assessment of Tobacco and Health Study Wave 1 (2013-2014). *International*
693 *journal of environmental research and public health* 2020;**17**.
- 694 11. Keith RJ, Fetterman JL, Orimoloye OA, Dardari Z, Lorkiewicz PK, Hamburg NM, DeFilippis
695 AP, Blaha MJ, Bhatnagar A. Characterization of Volatile Organic Compound Metabolites

- 696 in Cigarette Smokers, Electronic Nicotine Device Users, Dual Users, and Nonusers of
697 Tobacco. *Nicotine & tobacco research : official journal of the Society for Research on*
698 *Nicotine and Tobacco* 2020;**22**:264-272.
- 699 12. Atkinson R, Arey J. Atmospheric degradation of volatile organic compounds. *Chemical*
700 *reviews* 2003;**103**:4605-4638.
- 701 13. Li AJ, Pal VK, Kannan K. A review of environmental occurrence, toxicity, biotransformation
702 and biomonitoring of volatile organic compounds. *Environmental Chemistry and*
703 *Ecotoxicology* 2021.
- 704 14. Konkle SL, Zierold KM, Taylor KC, Riggs DW, Bhatnagar A. National secular trends in
705 ambient air volatile organic compound levels and biomarkers of exposure in the United
706 States. *Environ Res* 2020;**182**:108991.
- 707 15. Shuai J, Kim S, Ryu H, Park J, Lee CK, Kim GB, Ultra VU, Jr., Yang W. Health risk
708 assessment of volatile organic compounds exposure near Daegu dyeing industrial
709 complex in South Korea. *BMC public health* 2018;**18**:528.
- 710 16. Villeneuve PJ, Jerrett M, Su J, Burnett RT, Chen H, Brook J, Wheeler AJ, Cakmak S,
711 Goldberg MS. A cohort study of intra-urban variations in volatile organic compounds and
712 mortality, Toronto, Canada. *Environmental pollution (Barking, Essex : 1987)* 2013;**183**:30-
713 39.
- 714 17. DeJarnett N, Conklin DJ, Riggs DW, Myers JA, O'Toole TE, Hamzeh I, Wagner S, Chugh
715 A, Ramos KS, Srivastava S, Higdon D, Tollerud DJ, DeFilippis A, Becher C, Wyatt B,
716 McCracken J, Abplanalp W, Rai SN, Ciszewski T, Xie Z, Yeager R, Prabhu SD, Bhatnagar
717 A. Acrolein exposure is associated with increased cardiovascular disease risk. *J Am Heart*
718 *Assoc* 2014;**3**.
- 719 18. Abplanalp W, DeJarnett N, Riggs DW, Conklin DJ, McCracken JP, Srivastava S, Xie Z,
720 Rai S, Bhatnagar A, O'Toole TE. Benzene exposure is associated with cardiovascular
721 disease risk. *PLoS One* 2017;**12**:e0183602.
- 722 19. Ran J, Qiu H, Sun S, Yang A, Tian L. Are ambient volatile organic compounds
723 environmental stressors for heart failure? *Environmental Pollution* 2018;**242**:1810-1816.
- 724 20. Ye D, Klein M, Chang HH, Sarnat JA, Mulholland JA, Edgerton ES, Winquist A, Tolbert
725 PE, Sarnat SE. Estimating Acute Cardiorespiratory Effects of Ambient Volatile Organic
726 Compounds. *Epidemiology* 2017;**28**:197-206.
- 727 21. Matanoski GM, Tao XG. Styrene exposure and ischemic heart disease: a case-cohort
728 study. *American journal of epidemiology* 2003;**158**:988-995.
- 729 22. Lynch J, Jin L, Richardson A, Jagatheesan G, Lorkiewicz P, Xie Z, Theis WS, Shirk G,
730 Malovichko MV, Bhatnagar A, Srivastava S, Conklin DJ. Acute and chronic vascular
731 effects of inhaled crotonaldehyde in mice: Role of TRPA1. *Toxicol Appl Pharmacol*
732 2020;**402**:115120.
- 733 23. Jin L, Jagatheesan G, Guo L, Nystoriak M, Malovichko M, Lorkiewicz P, Bhatnagar A,
734 Srivastava S, Conklin DJ. Formaldehyde Induces Mesenteric Artery Relaxation via a

- 735 Sensitive Transient Receptor Potential Ankyrin-1 (TRPA1) and Endothelium-Dependent
736 Mechanism: Potential Role in Postprandial Hyperemia. *Front Physiol* 2019;**10**:277.
- 737 24. Awe SO, Adeagbo AS, D'Souza SE, Bhatnagar A, Conklin DJ. Acrolein induces
738 vasodilatation of rodent mesenteric bed via an EDHF-dependent mechanism. *Toxicol Appl*
739 *Pharmacol* 2006;**217**:266-276.
- 740 25. Conklin DJ, Malovichko MV, Zeller I, Das TP, Krivokhizhina TV, Lynch BH, Lorkiewicz P,
741 Agarwal A, Wickramasinghe N, Haberzettl P, Sithu SD, Shah J, O'Toole TE, Rai SN,
742 Bhatnagar A, Srivastava S. Biomarkers of Chronic Acrolein Inhalation Exposure in Mice:
743 Implications for Tobacco Product-Induced Toxicity. *Toxicol Sci* 2017;**158**:263-274.
- 744 26. Abplanalp WT, Wickramasinghe NS, Sithu SD, Conklin DJ, Xie Z, Bhatnagar A, Srivastava
745 S, O'Toole TE. Benzene Exposure Induces Insulin Resistance in Mice. *Toxicol Sci*
746 2019;**167**:426-437.
- 747 27. Sithu SD, Srivastava S, Siddiqui MA, Vladykovskaya E, Riggs DW, Conklin DJ, Haberzettl
748 P, O'Toole TE, Bhatnagar A, D'Souza SE. Exposure to acrolein by inhalation causes
749 platelet activation. *Toxicol Appl Pharmacol* 2010;**248**:100-110.
- 750 28. Conklin DJ, Barski OA, Lesgards JF, Juvan P, Rezen T, Rozman D, Prough RA,
751 Vladykovskaya E, Liu S, Srivastava S, Bhatnagar A. Acrolein consumption induces
752 systemic dyslipidemia and lipoprotein modification. *Toxicol Appl Pharmacol* 2010;**243**:1-
753 12.
- 754 29. Ismahil MA, Hamid T, Haberzettl P, Gu Y, Chandrasekar B, Srivastava S, Bhatnagar A,
755 Prabhu SD. Chronic oral exposure to the aldehyde pollutant acrolein induces dilated
756 cardiomyopathy. *Am J Physiol Heart Circ Physiol* 2011;**301**:H2050-2060.
- 757 30. Srivastava S, Sithu SD, Vladykovskaya E, Haberzettl P, Hoetker DJ, Siddiqui MA, Conklin
758 DJ, D'Souza SE, Bhatnagar A. Oral exposure to acrolein exacerbates atherosclerosis in
759 apoE-null mice. *Atherosclerosis* 2011;**215**:301-308.
- 760 31. Wang GW, Guo Y, Vondriska TM, Zhang J, Zhang S, Tsai LL, Zong NC, Bolli R, Bhatnagar
761 A, Prabhu SD. Acrolein consumption exacerbates myocardial ischemic injury and blocks
762 nitric oxide-induced PKCepsilon signaling and cardioprotection. *J Mol Cell Cardiol*
763 2008;**44**:1016-1022.
- 764 32. Heitzer T, Ylä-Herttuala S, Luoma J, Kurz S, Münzel T, Just Hr, Olschewski M, Drexler H.
765 Cigarette smoking potentiates endothelial dysfunction of forearm resistance vessels in
766 patients with hypercholesterolemia: role of oxidized LDL. *Circulation* 1996;**93**:1346-1353.
- 767 33. Pittilo M. Cigarette smoking, endothelial injury and cardiovascular disease. *International*
768 *Journal of Experimental Pathology* 2000;**81**:219-230.
- 769 34. Törnqvist H, Mills NL, Gonzalez M, Miller MR, Robinson SD, Megson IL, MacNee W,
770 Donaldson K, Söderberg S, Newby DE, Sandström T, Blomberg A. Persistent Endothelial
771 Dysfunction in Humans after Diesel Exhaust Inhalation. *American journal of respiratory*
772 *and critical care medicine* 2007;**176**:395-400.

- 773 35. McGraw KE, Riggs DW, Rai S, Navas-Acien A, Xie Z, Lorkiewicz P, Lynch J, Zafar N,
774 Krishnasamy S, Taylor KC, Conklin DJ, DeFilippis AP, Srivastava S, Bhatnagar A.
775 Exposure to volatile organic compounds – acrolein, 1,3-butadiene, and crotonaldehyde –
776 is associated with vascular dysfunction. *Environmental research* 2021;**196**:110903.
- 777 36. Hill JM, Zalos G, Halcox JP, Schenke WH, Waclawiw MA, Quyyumi AA, Finkel T.
778 Circulating endothelial progenitor cells, vascular function, and cardiovascular risk. *New*
779 *England Journal of Medicine* 2003;**348**:593-600.
- 780 37. Zafar N, Krishnasamy SS, Shah J, Rai SN, Riggs DW, Bhatnagar A, O'Toole TE.
781 Circulating angiogenic stem cells in type 2 diabetes are associated with glycemic control
782 and endothelial dysfunction. *PloS one* 2018;**13**:e0205851.
- 783 38. Werner N, Nickenig G. Influence of cardiovascular risk factors on endothelial progenitor
784 cells: limitations for therapy? *Arteriosclerosis, thrombosis, and vascular biology*
785 2006;**26**:257-266.
- 786 39. Werner N, Kosiol S, Schiegl T, Ahlers P, Walenta K, Link A, Böhm M, Nickenig G.
787 Circulating endothelial progenitor cells and cardiovascular outcomes. *New England*
788 *Journal of Medicine* 2005;**353**:999-1007.
- 789 40. Yeager R, Riggs DW, DeJarnett N, Srivastava S, Lorkiewicz P, Xie Z, Krivokhizhina T,
790 Keith RJ, Srivastava S, Browning M, Zafar N, Krishnasamy S, DeFilippis A, Turner J, Rai
791 SN, Bhatnagar A. Association between residential greenness and exposure to volatile
792 organic compounds. *The Science of the total environment* 2020;**707**:135435.
- 793 41. O'Toole TE, Hellmann J, Wheat L, Haberzettl P, Lee J, Conklin DJ, Bhatnagar A, Pope
794 CA, 3rd. Episodic exposure to fine particulate air pollution decreases circulating levels of
795 endothelial progenitor cells. *Circ Res* 2010;**107**:200-203.
- 796 42. Fadini GP, de Kreutzenberg S, Albiero M, Coracina A, Pagnin E, Baesso I, Cignarella A,
797 Bolego C, Plebani M, Nardelli GB. Gender differences in endothelial progenitor cells and
798 cardiovascular risk profile: the role of female estrogens. *Arteriosclerosis, thrombosis, and*
799 *vascular biology* 2008;**28**:997-1004.
- 800 43. Xia WH, Yang Z, Xu SY, Chen L, Zhang XY, Li J, Liu X, Qiu YX, Shuai XT, Tao J. Age-
801 Related Decline in Reendothelialization Capacity of Human Endothelial Progenitor Cells
802 Is Restored by Shear Stress. *Hypertension* 2012;**59**:1225-1231.
- 803 44. Samman Tahhan A, Hammadah M, Kelli HM, Kim JH, Sandesara PB, Alkhoder A, Kaseer
804 B, Gafeer MM, Topel M, Hayek SS, O'Neal WT, Obideen M, Ko Y-A, Liu C, Hesaroieh I,
805 Mahar E, Vaccarino V, Waller EK, Quyyumi AA. Circulating Progenitor Cells and Racial
806 Differences. *Circulation research* 2018;**123**:467-476.
- 807 45. Tobler K, Freudenthaler A, Baumgartner-Parzer SM, Wolzt M, Ludvik B, Nansalmaa E,
808 Nowotny PJ, Seidinger D, Steiner S, Luger A, Artwohl M. Reduction of both number and
809 proliferative activity of human endothelial progenitor cells in obesity. *International Journal*
810 *of Obesity* 2010;**34**:687-700.
- 811 46. Fadini GP, Sartore S, Agostini C, Avogaro A. Significance of Endothelial Progenitor Cells
812 in Subjects With Diabetes. *Diabetes care* 2007;**30**:1305-1313.

- 813 47. Giannotti G, Doerries C, Mocharla Pavani S, Mueller Maja F, Bahlmann Ferdinand H,
814 Horvath T, Jiang H, Sorrentino Sajoscha A, Steenken N, Manes C, Marzilli M, Rudolph
815 KL, Lüscher Thomas F, Drexler H, Landmesser U. Impaired Endothelial Repair Capacity
816 of Early Endothelial Progenitor Cells in Prehypertension. *Hypertension* 2010;**55**:1389-
817 1397.
- 818 48. Park SK, Tao Y, Meeker JD, Harlow SD, Mukherjee B. Environmental risk score as a new
819 tool to examine multi-pollutants in epidemiologic research: an example from the NHANES
820 study using serum lipid levels. *PloS one* 2014;**9**:e98632.
- 821 49. Bobb JF, Valeri L, Claus Henn B, Christiani DC, Wright RO, Mazumdar M, Godleski JJ,
822 Coull BA. Bayesian kernel machine regression for estimating the health effects of multi-
823 pollutant mixtures. *Biostatistics* 2015;**16**:493-508.
- 824 50. Bobb JF, Henn BC, Valeri L, Coull BA. Statistical software for analyzing the health effects
825 of multiple concurrent exposures via Bayesian kernel machine regression. *Environmental*
826 *Health* 2018;**17**:1-10.
- 827 51. Ran J, Qiu H, Sun S, Tian L. Short-term effects of ambient benzene and TEX (toluene,
828 ethylbenzene, and xylene combined) on cardiorespiratory mortality in Hong Kong. *Environ*
829 *Int* 2018;**117**:91-98.
- 830 52. Xu X, Freeman NC, Dailey AB, Ilacqua VA, Kearney GD, Talbott EO. Association between
831 exposure to alkylbenzenes and cardiovascular disease among National Health and
832 Nutrition Examination Survey (NHANES) participants. *Int J Occup Environ Health*
833 2009;**15**:385-391.
- 834 53. Shim YH, Ock JW, Kim Y-J, Kim Y, Kim SY, Kang D. Association between Heavy Metals,
835 Bisphenol A, Volatile Organic Compounds and Phthalates and Metabolic Syndrome.
836 *International journal of environmental research and public health* 2019;**16**:671.
- 837 54. Penn A, Snyder CA. 1,3-Butadiene exposure and cardiovascular disease. *Mutat Res*
838 2007;**621**:42-49.
- 839 55. Capella KM, Roland K, Geldner N, Rey deCastro B, De Jesús VR, van Bommel D, Blount
840 BC. Ethylbenzene and styrene exposure in the United States based on urinary mandelic
841 acid and phenylglyoxylic acid: NHANES 2005–2006 and 2011–2012. *Environmental*
842 *research* 2019;**171**:101-110.
- 843 56. Mori TA, Burke V, Puddey IB, Watts GF, O'Neal DN, Best JD, Beilin LJ. Purified
844 eicosapentaenoic and docosahexaenoic acids have differential effects on serum lipids and
845 lipoproteins, LDL particle size, glucose, and insulin in mildly hyperlipidemic men. *The*
846 *American journal of clinical nutrition* 2000;**71**:1085-1094.
- 847 57. Tang W, Hemm I, Eisenbrand G. Estimation of human exposure to styrene and
848 ethylbenzene. *Toxicology* 2000;**144**:39-50.
- 849 58. Vasa M, Fichtlscherer S, Aicher A, Adler K, Urbich C, Martin H, Zeiher AM, Dimmeler S.
850 Number and migratory activity of circulating endothelial progenitor cells inversely correlate
851 with risk factors for coronary artery disease. *Circulation research* 2001;**89**:e1-e7.

- 852 59. Imanishi T, Moriwaki C, Hano T, Nishio I. Endothelial progenitor cell senescence is
853 accelerated in both experimental hypertensive rats and patients with essential
854 hypertension. *Journal of hypertension* 2005;**23**:1831-1837.
- 855 60. Li MJ, Zeng T. The deleterious effects of N,N-dimethylformamide on liver: A mini-review.
856 *Chemico-biological interactions* 2019;**298**:129-136.
- 857 61. Haberzettl P, Lee J, Duggineni D, McCracken J, Bolanowski D, O'Toole TE, Bhatnagar A,
858 Conklin DJ. Exposure to ambient air fine particulate matter prevents VEGF-induced
859 mobilization of endothelial progenitor cells from the bone marrow. *Environmental health*
860 *perspectives* 2012;**120**:848.
- 861 62. Liberda EN, Cuevas AK, Gillespie PA, Grunig G, Qu Q, Chen LC. Exposure to inhaled
862 nickel nanoparticles causes a reduction in number and function of bone marrow
863 endothelial progenitor cells. *Inhalation toxicology* 2010;**22**:95-99.
- 864 63. Brook RD, Bard RL, Kaplan MJ, Yalavarthi S, Morishita M, Dvonch JT, Wang L, Yang H-
865 y, Spino C, Mukherjee B. The effect of acute exposure to coarse particulate matter air
866 pollution in a rural location on circulating endothelial progenitor cells: results from a
867 randomized controlled study. *Inhalation toxicology* 2013;**25**:587-592.
- 868 64. DeJarnett N, Yeager R, Conklin DJ, Lee J, O'Toole TE, McCracken J, Abplanalp W,
869 Srivastava S, Riggs DW, Hamzeh I, Wagner S, Chugh A, DeFilippis A, Ciszewski T, Wyatt
870 B, Becher C, Higdon D, Ramos KS, Tollerud DJ, Myers JA, Rai SN, Shah J, Zafar N,
871 Krishnasamy SS, Prabhu SD, Bhatnagar A. Residential Proximity to Major Roadways Is
872 Associated With Increased Levels of AC133+ Circulating Angiogenic Cells. *Arterioscler*
873 *Thromb Vasc Biol* 2015.
- 874 65. Yeager R, Riggs DW, DeJarnett N, Tollerud DJ, Wilson J, Conklin DJ, O'Toole TE,
875 McCracken J, Lorkiewicz P, Xie Z, Zafar N, Krishnasamy SS, Srivastava S, Finch J, Keith
876 RJ, DeFilippis A, Rai SN, Liu G, Bhatnagar A. Association Between Residential
877 Greenness and Cardiovascular Disease Risk. *J Am Heart Assoc* 2018;**7**:e009117.
- 878 66. Chotinantakul K, Leraanansaksiri W. Hematopoietic Stem Cell Development, Niches, and
879 Signaling Pathways. *Bone Marrow Research* 2012;**2012**:16.
- 880 67. Boos CJ, Lip GYH, Blann AD. Circulating Endothelial Cells in Cardiovascular Disease.
881 *Journal of the American College of Cardiology* 2006;**48**:1538-1547.
- 882 68. Qian X, Wan Y, Wang A, Xia W, Yang Z, He Z, Xu S. Urinary metabolites of multiple
883 volatile organic compounds among general population in Wuhan, central China: Inter-day
884 reproducibility, seasonal difference, and their associations with oxidative stress
885 biomarkers. *Environmental Pollution* 2021;**289**:117913.
- 886 69. Song W, Han Q, Wan Y, Qian X, Wei M, Jiang Y, Wang Q. Repeated measurements of
887 21 urinary metabolites of volatile organic compounds and their associations with three
888 selected oxidative stress biomarkers in 0– 7-year-old healthy children from south and
889 central China. *Chemosphere* 2021:132065.

890 70. Mobarrez F, Antoniewicz L, Bosson JA, Kuhl J, Pisetsky DS, Lundbäck M. The effects of
891 smoking on levels of endothelial progenitor cells and microparticles in the blood of healthy
892 volunteers. *PloS one* 2014;**9**:e90314-e90314.

893

894

895 **Figure Legends**

896 **Scheme 1. Recruitment Flow Chart.**

Figure 1 Association between the urinary metabolites of VOCs and circulating levels of CACs and heat map of p-values for pair-wise associations between the urinary metabolites of VOCs and circulating levels of CACs. (A) False Discovery rate adjusted p-values for total sample (n=603). (B) False Discovery rate adjusted p-values for non-smokers (n=375). Black grids represent positive associations and blue grids represent negative associations. The intensity of color represents p-values, with darker grids indicating smaller p-values. (C) The X axis represents % difference (\pm 95% confidence interval) per 2-fold increase in VOC metabolite for each CAC sub-population (CAC 1-15) in non-smokers (n=375). Generalized linear models were adjusted for age, sex, race, BMI, diabetes, hypertension, and daily PM_{2.5} levels. Models for the full sample were additionally adjusted for smoking status. Abbreviations: CEMA, N-Acetyl-S-(2-carboxyethyl)-L-cysteine; 3-HPMA, N-Acetyl-S-(3-hydroxypropyl)-L-cysteine; AAMA, N-Acetyl-S-(2-carbamoylethyl)-L-cysteine; CYMA, N-Acetyl-S-(2-cyanoethyl)-L-cysteine; MU, *trans, trans*-muconic acid; DHBMA, N-Acetyl-S-(3,4-dihydroxybutyl)-L-cysteine; MHBMA3, N-Acetyl-S-(4-hydroxy-2-buten-1-yl)-L-cysteine; HPMMA, N-Acetyl-S-(3-hydroxypropyl-1-methyl)-L-cysteine; AMCC, N-Acetyl-S-(N-methylcarbamoyl)-L-cysteine; PGA, Phenylglyoxylic acid ; 2HPMA, N-Acetyl-S-(2-hydroxypropyl)-L-cysteine; PHEMA, N-Acetyl-S-(1-phenyl-2-hydroxyethyl)-L-cysteine +; MA, Mandelic Acid; BMA, N-Acetyl-S-(benzyl)-L-cysteine; 2MHA, 2-Methylhippuric acid; 3MHA+4MHA, 3-Methylhippuric acid + 4-Methylhippuric acid. * Represents significant association.

897

898 **Figure 2.** Association between Environmental Risk Scores and Circulating Angiogenic Cells in
899 non-smokers. Environmental Risk score was calculated for each cell population as follows: 1.
900 The urinary levels of VOC metabolites were log transformed. 2. Each log transformed VOC
901 metabolite level was regressed against the levels of the cells separately, while adjusting for
902 covariates. 3. ERS was calculated as the sum of the log transformed VOC metabolite levels x
903 (regression coefficient/standard error). Mean and standard deviation of the ERS was calculated
904 for each CAC. Values are mean \pm SD of environmental risk score for each cell population.

905

906 **Figure 3.** Exposure-response relationships between selected CACs and VOCs in non-smokers.
907 Non-linear relationships between MU metabolites and CACs (**A**); VOCs and CAC-4 (**B**); and
908 VOCs and CAC-8 (**C**) were examined using restricted cubic splines. Tests for non-linearity used
909 the likelihood ratio test. Models are adjusted for age, sex, race, BMI, diabetes, hypertension, beta-
910 blocker use, and daily PM_{2.5} levels. P-value represents likelihood ratio test for non-linearity.

911

912 **Figure 4.** Association between circulating angiogenic cells and multiple VOCs in the same model
913 (non-smokers). Values represent % difference per 2-fold increase in VOC metabolite. VOC
914 metabolites are ordered by strength of their regression coefficients. Only significant metabolites
915 from one-pollutant models were entered in the model. Models adjusted for age, sex, race, BMI,
916 diabetes, hypertension, and daily PM_{2.5} levels.

917

918 **Figure 5.** Overall association of the VOC metabolite mixture (and 95% credible intervals) with
919 selected CACs, when all exposures are set to a given percentile relative to median. Models (non-
920 smokers; n=375) were adjusted for age, sex, race, BMI, diabetes, hypertension, and daily PM_{2.5}
921 levels.

922

923 **Figure 6.** Subgroup analysis of the relationship between different CAC population with the urinary
924 metabolites of benzene (MU), 1,3-butadiene (MHBMA3), ethylbenzene/styrene (PGA), and N,N-
925 Dimethylformamide (AMCC) in non-smokers. Generalized linear models were adjusted for
926 covariates - age, sex, race, BMI, diabetes, hypertension, and daily PM_{2.5} levels, unless the
927 covariates were used for stratification. *Significant interaction between subgroups in the full
928 model.

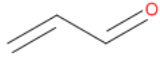
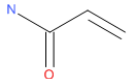


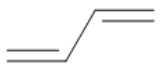
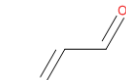

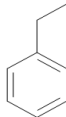

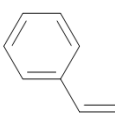
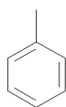

929

930 **Scheme 2.** Gating Strategy for the Human Circulating Angiogenic Cells. Blue indicates negative
931 environmental risk score; orange indicates positive environmental risk score.

932

Supplemental Results

Supplemental Table 1. Urinary VOC metabolites and Parent Compound

VOC Structure	VOC	VOC metabolite	Common abbr.	LOD (ng/ml)
	Acrolein	N-Acetyl-S-(2-carboxyethyl)-L-cysteine	CEMA	20.35
		N-Acetyl-S-(3-hydroxypropyl)-L-cysteine	3HPMA	9.12
	Acrylamide	N-Acetyl-S-(2-carbamoyl)-L-cysteine	AAMA	3.03
	Acrylonitrile	N-Acetyl-S-(2-cyanoethyl)-L-cysteine	CYMA	1.25
	Benzene	<i>trans, trans</i> -muconic acid	MU	6.41
	1,3-Butadiene	N-Acetyl-S-(3,4-dihydroxybutyl)-L-cysteine	DHBMA	6.76
		N-Acetyl-S-(4-hydroxy-2-buten-1-yl)-L-cysteine	MHBMA3	1.44
	Crotonaldehyde	N-Acetyl-S-(3-hydroxypropyl-1-methyl)-L-cysteine	HPMMA	1.11
	N,N-Dimethylformamide	N-Acetyl-S-(N-methylcarbamoyl)-L-cysteine	AMCC	9.87
	Ethylbenzene, styrene	Phenylglyoxylic acid	PGA	6.86
	Propylene oxide	N-Acetyl-S-(2-hydroxypropyl)-L-cysteine	2HPMA	3.97
	Styrene, ethylbenzene	N-Acetyl-S-(1-phenyl-2-hydroxyethyl)-L-cysteine +	PHEMA	1.21
		N-Acetyl-S-(2-phenyl-2-hydroxyethyl)-L-cysteine		
		Mandelic Acid	MA	12.93
	Toluene	N-Acetyl-S-(benzyl)-L-cysteine	BMA	0.45
	Xylene	2-Methylhippuric acid	2MHA	1.24
		3-Methylhippuric acid + 4-Methylhippuric acid	3MHA+4MHA	6.85

Supplemental Table 2. Summary Statistics of urinary VOC metabolites in non-smokers for specific subgroups.

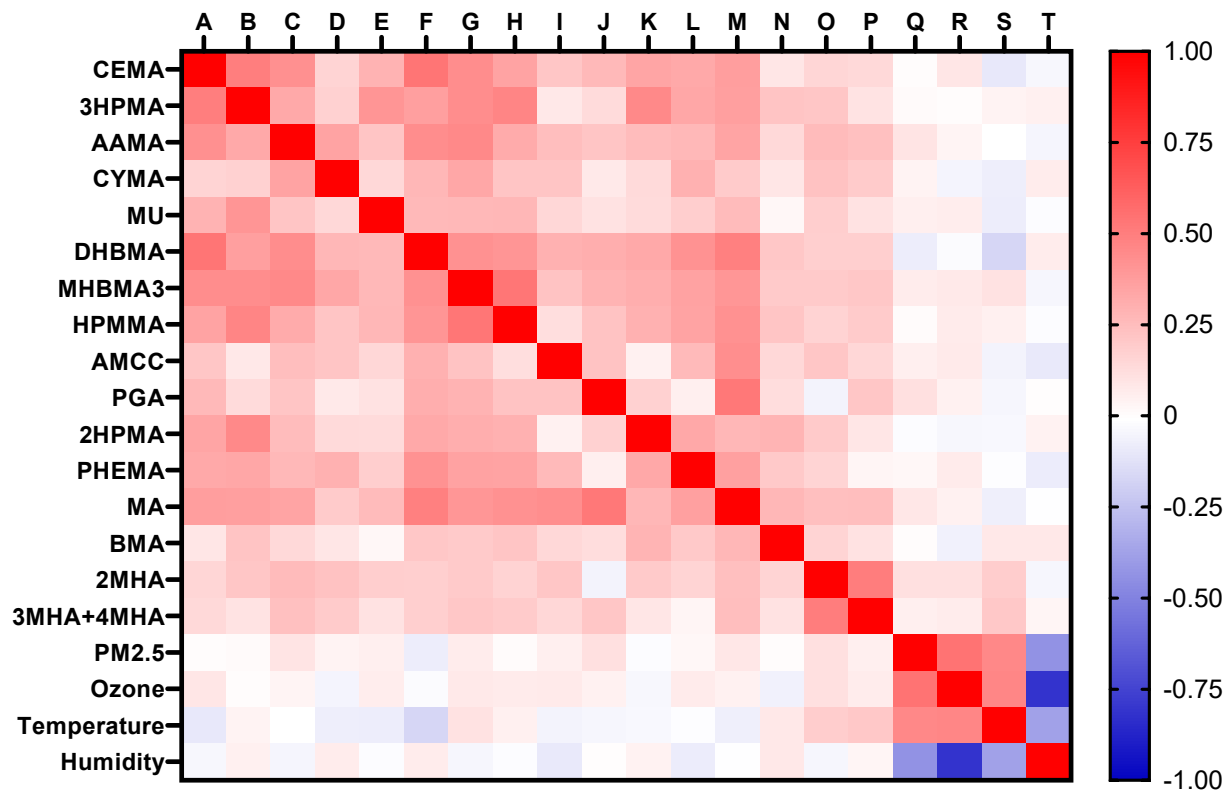
VOC metabolite	Male	Female	White	Black	Age <55	Age ≥55	BMI <30	BMI ≥30
n	135	240	311	43	202	173	199	176
CEMA	104.1 ± 87.6	122.4 ± 130.4	111.2 ± 102.5	136.1 ± 110.7	102.1 ± 78.4	131.8 ± 148.7	113.1 ± 132.6	118.9 ± 96.8
3HPMA	326.2 ± 269.2	350.2 ± 380.8	333.6 ± 313.3	330.3 ± 395.3	337.3 ± 325.8	346.5 ± 366.3	346.3 ± 310.9	336.1 ± 380.0
AAMA	73.4 ± 56.9	71.8 ± 64.2	71.2 ± 63.5	83.2 ± 53.5	72.2 ± 56.8	72.7 ± 66.9	77.9 ± 72.1	66.2 ± 46.3
CYMA	4.8 ± 12.3	6.1 ± 19.4	4.7 ± 15.5	12 ± 23.3	6.4 ± 17.0	4.8 ± 17.4	6.8 ± 20.6	4.4 ± 12.2
MU	125.8 ± 222.5	137.7 ± 209.4	146 ± 228.7	62.2 ± 52.1	126.0 ± 208.3	142.0 ± 220.8	134.8 ± 229.1	131.7 ± 196.2
DHBMA	348.7 ± 182.2	356.4 ± 172.6	363.6 ± 181.3	311.8 ± 148.3	328.0 ± 157.2	383.5 ± 191.6	355.0 ± 166.6	352.1 ± 186.3
MHBMA3	8.4 ± 8.6	9.5 ± 8	9.3 ± 8.6	8.2 ± 6.1	9.2 ± 7.9	9.0 ± 8.6	10.2 ± 9.6	7.9 ± 6.1
HPMMA	286.1 ± 271.3	321.5 ± 324.1	305.1 ± 265.4	347.3 ± 529.9	275.4 ± 241.8	328.6 ± 267.7	315.1 ± 292.7	282.8 ± 203.8
AMCC	140 ± 99.4	171.9 ± 125.1	169.3 ± 121.1	118.5 ± 74.2	143.2 ± 104.3	180.5 ± 128.3	165.7 ± 119.9	154.5 ± 114.4
PGA	202.2 ± 117.2	239.1 ± 157.9	232.2 ± 151.8	196.3 ± 105	215.5 ± 149.0	238.0 ± 140.9	226.8 ± 154.1	224.9 ± 135.7
2HPMA	67.7 ± 160.1	53.1 ± 95.8	60.1 ± 131.4	40.8 ± 31.8	58.1 ± 116.8	58.6 ± 130.0	67.1 ± 147.8	48.4 ± 86.0
PHEMA	2.3 ± 5.4	2.2 ± 3	2.3 ± 4.4	1.7 ± 1.2	2.1 ± 3.9	2.3 ± 4.2	2.2 ± 3.9	2.2 ± 4.2
MA	206.3 ± 124.1	247.7 ± 183.3	243.7 ± 175.6	177.8 ± 91.8	210.7 ± 139.5	258.7 ± 188.7	231.4 ± 156.0	234.5 ± 176.0
BMA	10.7 ± 12.4	12 ± 11.9	11.4 ± 12.1	11.6 ± 11.8	10.5 ± 11.9	12.7 ± 12.2	12.2 ± 12.8	10.7 ± 11.1
2MHA	26.3 ± 37.1	36.5 ± 89.8	33.8 ± 78.7	25.5 ± 32.6	27.6 ± 50.4	32.9 ± 56.7	37.9 ± 63.7	21.2 ± 36.8
3MHA+4MHA	189 ± 313.1	220.5 ± 454.3	213 ± 436	193.8 ± 205.9	179.8 ± 310.6	212.5 ± 293.9	227.4 ± 367.2	158.4 ± 204.2

Values represent Mean ± SD. Units: µg/g creatinine. Bold values represent significant difference between specific subgroups, based on T-test of log-transformed VOC metabolites.

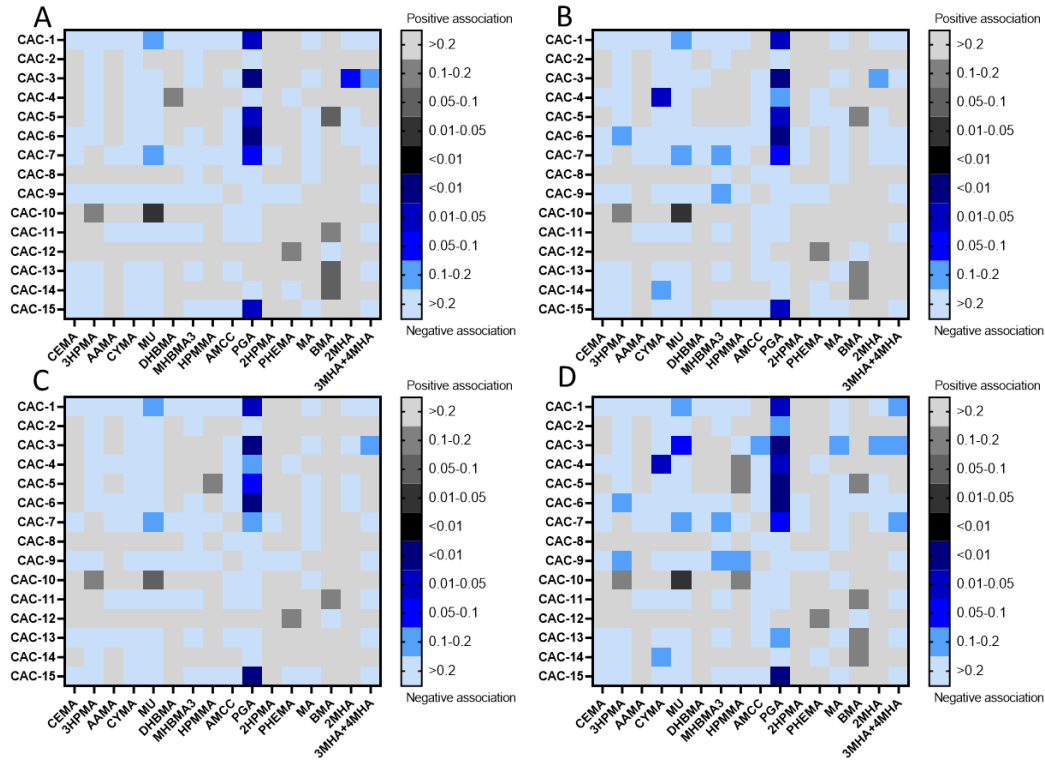
Supplemental Table 2 continued. Summary Statistics of urinary VOC metabolites in non-smokers for specific subgroups.

VOC metabolite	Non-Diabetic	Diabetic	Normotensive	Hypertensive	Low PM _{2.5}	High PM _{2.5}
n	310	65	219	156	189	186
CEMA	116.8 ± 123.3	111.2 ± 81	104.7 ± 85.1	131.4 ± 149.8	117.0 ± 92.4	114.6 ± 137.9
3HPMA	338.7 ± 341.3	355 ± 362.4	327.9 ± 315.8	360.6 ± 381.7	371.4 ± 411.7	311.2 ± 257.0
AAMA	74.6 ± 64.3	62.2 ± 45.5	70.7 ± 55.6	74.9 ± 69.3	65.0 ± 41.6	80.0 ± 76.1
CYMA	6.1 ± 18.5	3.6 ± 8.7	5.7 ± 15.9	5.6 ± 18.9	4.4 ± 12.3	7.0 ± 21.0
MU	128.1 ± 213.7	158.4 ± 215.3	127.5 ± 227.6	141.6 ± 193.7	142.7 ± 257.3	123.9 ± 158.5
DHBMA	354.3 ± 172.8	350.4 ± 191.2	347.4 ± 164.4	362.4 ± 191.2	368.6 ± 182.0	338.5 ± 168.6
MHBMA3	9.3 ± 8.6	8.3 ± 6.2	8.9 ± 7.7	9.5 ± 8.9	9.1 ± 8.4	9.2 ± 8.1
HPMMA	300.9 ± 260.8	346 ± 466.9	277.8 ± 228	352.1 ± 387.2	308.9 ± 278.3	290.9 ± 229.5
AMCC	157.3 ± 108.6	175.1 ± 152.4	154.8 ± 109.2	168.3 ± 127.8	153.5 ± 102.8	167.4 ± 130.4
PGA	219.3 ± 143.1	257.2 ± 154	222.7 ± 156.8	230.3 ± 128.6	208.3 ± 133.0	243.7 ± 155.6
2HPMA	62 ± 132.5	40.9 ± 55.8	66.1 ± 137	47.4 ± 99.3	64.8 ± 144.2	51.8 ± 96.5
PHEMA	2.3 ± 4.4	1.8 ± 1.9	2.3 ± 4.6	2.1 ± 3.2	2.1 ± 3.6	2.3 ± 4.5
MA	228.5 ± 155.5	253.4 ± 206.9	223.9 ± 154.4	245.3 ± 179.6	217.1 ± 146.2	248.9 ± 182.1
BMA	12 ± 12.7	9.4 ± 8.1	11.7 ± 12.7	11.3 ± 11.1	10.8 ± 11.4	12.3 ± 12.7
2MHA	35 ± 80	22.5 ± 46.6	37.4 ± 91.6	26.5 ± 43.2	32.8 ± 64.7	27.3 ± 38.8
3MHA+4MHA	223.7 ± 445.8	140.2 ± 115.9	226.4 ± 480	185.2 ± 282.4	211.2 ± 368.4	178.4 ± 218.4

Values represent Mean ± SD. Units: µg/g creatinine. Bold values represent significant difference between specific subgroups, based on T-test of log-transformed VOC metabolites. Low PM_{2.5} ranges from 1.2-7.8 µg/m³. High PM_{2.5} ranges from 8.0-17.7 µg/m³.



Supplemental Figure 1. Pearson correlation coefficients between log-transformed urinary VOC metabolites and ambient pollutant and meteorological data (non-smokers; n=375).



Supplemental Figure 2. Heat map of pair-wise associations between the urinary metabolites of VOCs and circulating levels of CACs. False Discovery rate adjusted p-values. Generalized linear models (non-smokers; n=375) were adjusted for age, sex, race, BMI, diabetes, hypertension and: daily PM_{2.5} + Temperature (A); daily PM_{2.5} + Humidity (B); Ozone + Temperature (C); Ozone + Humidity (D). Black grids represent positive associations and blue grids represent negative associations. The intensity of colors represent p-values, with darker grids indicating smaller p-values.