Environmental Exposure to Volatile Organic Compounds is Associated with Endothelial Injury

*Daniel W. Riggs,[1-5] *Marina V. Malovichko,[1-4] Hong Gao,[1-4] Katlyn E. McGraw,[1-4] Breandon S. Taylor,[1,2] Tatiana Krivokhizhina,[1-4] Shesh N. Rai,[1,2,6] Rachel J. Keith,[1,2,4] Aruni Bhatnagar,[1-4] and Sanjay Srivastava[1-4]

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

*These people contributed equally to the manuscript

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

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

ABSTRACT

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

Approach and Results: In this cross-sectional study, we recruited 375 non-smokers with low-to-high CVD risk and measured 15 subpopulations of CACs by flow cytometry and 16 urinary metabolites of 12 VOCs by LC/MS/MS. Associations between CAC and VOC metabolite levels were examined using generalized linear models. Our data show that VOC exposure varied by sex, age, and race. In single pollutant models, metabolites of xylene, benzene, and 1,3-butadiene were negatively associated with CAC levels. The metabolite of acrylonitrile was negatively associated with $CD45^{dim}/CD146^{+}/CD34^{+}/AC133^{+}$ cells and $CD45^{+}/CD146^{+}/AC133^{+}$, and the toluene metabolite with $AC133^{+}$ cells. Multipollutant models showed a negative association with metabolites of ethylbenzene/styrene, benzene, and xylene with $CD45^{dim}/CD146^{+}/CD34^{+}$ cells, independent of other VOC metabolite levels. Cumulative VOC risk score showed a strong negative association with $CD45^{dim}/CD146^{+}/CD34^{+}$ cells, suggesting that total VOC exposure has a cumulative effect on pro-angiogenic cells. Dose-response relationship showed an increase in CAC levels at low, but depletion at higher levels of VOC exposure. Sex and race, hypertension, and diabetes significantly modified VOC associated CAC depletion.

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

Keywords: Circulating angiogenic cells, VOCs, endothelial injury, benzene, ethylbenzene, xylene

The study was approved by the University of Louisville Institutional Review Board (IRB 15.1260).

1. INTRODUCTION

Exposure to ambient air pollution has been widely recognized to be an important contributor to premature mortality and morbidity. The Global Burden of Disease (GBD) study estimates that in 2015, 4.2 million deaths could be attributed to ambient air pollutants, with an additional 2.8 million deaths due to household air pollution (Cohen et al., 2017). Exposure to fine particulate matter in the ambient air, defined as particle size < 2.5 μm in diameter ($PM_{2.5}$), is associated with an increased risk for hypertension (Yang et al., 2018), myocardial infarction (Mustafić et al., 2012), type 2 diabetes (Brook et al., 2013b; Rajagopalan and Brook, 2012), stroke (Shah et al., 2015), and cardiovascular mortality (Rajagopalan et al., 2018). The emphasis on $PM_{2.5}$ and cardiovascular disease (CVD) has been motivated largely by the availability of historical monitoring data, however in an urban setting, over 90% of pollutant mass consists of gases or vaporous compounds, such as volatile organic compounds (VOCs) (Brook et al., 2010). In comparison with $PM_{2.5}$, data on VOCs tend to exhibit higher temporal and spatial variability.

Most studies on VOC have focused on tobacco smoke (De Jesús et al., 2020; Keith et al., 2020; Lorkiewicz et al., 2019) which represents a major source of human exposures. However, non-tobacco derived VOCs are ubiquitous in urban environments. High concentrations of VOCs such as alkanes, alkenes, aromatic hydrocarbons, and aldehydes, are emitted into the atmosphere from both anthropogenic and biogenic sources, and although biogenic VOCs are more abundant worldwide, in urban areas, anthropogenic VOCs are often higher (Atkinson and Arey, 2003). Outdoor VOC exposures can originate from vehicle emissions, industrial facilities, landfills and

89 hazardous waste sites, while indoor exposure can originate from diverse sources, such
90 as paints, cleaning agents, personal care products, and solvents (Li et al., 2021). Our
91 analysis of the secular trends of VOC exposures in the US suggests that even though
92 there was a decline in the ambient air VOC levels, individual-level exposure increased
93 from 2005-2014 (Konkle et al., 2020). Hence, there is urgent need to assess the health
94 impacts of VOC exposures.

95 VOC exposure have been associated with a wide range of acute and chronic
96 health effects, including allergies, respiratory diseases, liver dysfunction, and cancer
97 (Shuai et al., 2018; Villeneuve et al., 2013). Moreover, exposures to ambient levels of
98 acrolein, benzene, and xylene have been associated with elevated CVD risk (Abplanalp
99 et al., 2017; DeJarnett et al., 2014). In an intra-urban analysis of VOC gradients, higher
100 levels of VOCs were found to be associated with increased cardiovascular and cancer
101 mortality (Villeneuve et al., 2013). Moreover, longitudinal time-series studies have shown
102 that ambient VOCs, such as alkenes and alkynes, are associated with increased heart
103 failure emergency hospitalizations (Ran et al., 2018b), as well as overall cardiovascular
104 hospitalizations (Ye et al., 2017), and that occupational exposure to VOC such as styrene
105 increases the relative risk of death from ischemic heart disease (Matanoski and Tao,
106 2003).

107 The biological plausibility that VOCs can induce cardiovascular injury and
108 exacerbate CVD is extensively supported by a range of animal studies showing that
109 exposures to benzene, acrolein, and formaldehyde, can lead to changes in blood
110 pressure regulation (Awe et al., 2006; Jin et al., 2019; Lynch et al., 2020), depletion of
111 circulating stem cells (Conklin et al., 2017), insulin resistance (Abplanalp et al., 2019),

112 and platelet activation (Sithu et al., 2010). Exposure to these VOCs also induces systemic
113 dyslipidemia, lipoprotein modification (Conklin et al., 2010), dilated cardiomyopathy
114 (Ismahil et al., 2011); and lead to the acceleration of atherogenesis (Srivastava et al.,
115 2011) and exacerbation of myocardial ischemic injury (Wang et al., 2008). Nonetheless,
116 despite this evidence, little is known about the effects of VOCs on the mechanisms and
117 processes that contribute to cardiovascular disease in humans. Previous work has shown
118 that in humans, high levels of VOC exposures from sources such as cigarette smoke
119 (Heitzer et al., 1996; Pittilo, 2000) and diesel exhaust (Törnqvist et al., 2007) are
120 associated with endothelial dysfunction. We have previously found that exposure to
121 acrolein and 1,3-butadiene are associated with lower reactive hyperemia index, a
122 measure of endothelial function (Li et al., 2021; McGraw et al., 2021). The endothelium
123 is a first line of defense of cardiovascular tissues against environmental toxins, and
124 endothelial function is a sensitive reflection of vascular and systemic health. Therefore,
125 to identify and understand the effects of VOC exposures and cardiovascular health, we
126 measured circulating angiogenic cells (CAC), also referred to as endothelial progenitor
127 cells (EPC). It is believed that these cell aid in the repair of the endothelium after injury,
128 and are good biomarkers of vascular function (Hill et al., 2003; Zafar et al., 2018). CAC
129 levels are associated with several CVD risk factors (Werner and Nickenig, 2006), and are
130 predictive of both cardiovascular events and cardiovascular mortality (Werner et al.,
131 2005). Thus, using CACs as surrogate biomarkers of endothelial health, we examined
132 their association with urinary levels of VOC metabolite levels in non-smokers with low to
133 severe CVD risk.

134 **2. METHODS**

135 **2.1 Study Population and Design**

136 In a cross-sectional study, adult study participants, aged 25-70 years, were
137 recruited from the Louisville, KY area. Recruitment occurred in the summer months (May
138 to September) in the years 2018 and 2019, to minimize seasonal outdoor VOC
139 heterogeneity. Persons unwilling or unable to provide informed consent or with significant
140 and/or severe comorbidities were excluded. These included subjects that had hepatitis,
141 HIV/AIDS, active treatment for cancer, and active bleeding including wounds. Other
142 exclusions included subjects with body weight less than 100 pounds, pregnant women,
143 prisoners, and other vulnerable populations, those with conditions known to effect
144 peripheral blood cell counts and bone marrow function.

145 At the time of enrollment non-fasting blood and spot urine were collected from
146 patients who met the enrollment criteria and gave written consent. Demographic
147 information, including age, sex, ethnicity, residential address, smoking status history and
148 secondhand smoke exposure (verified by measurement of urinary cotinine
149 concentrations), and BMI (calculated from self-reported height and weight) was obtained
150 by using a questionnaire. Information on medication usage, blood pressure, and CVD
151 history, including incidence of heart attack, heart failure, angina, hypertension,
152 hypercholesterolemia, diabetes, obesity, stroke, and revascularization were also obtained
153 through the questionnaire. For the current study, hypertension was defined as a previous
154 diagnosis of hypertension or systolic blood pressure greater than 140 mmHg. Diabetes
155 was defined as previous diagnosis of diabetes or HbA1C greater than 6.5. The study was
156 approved by the University of Louisville Institutional Review Board (IRB 15.1260). A total
157 of 735 individuals met the inclusion criteria (Scheme 1).

158 Of those enrolled in the study, 115 individuals did not provide blood samples, or did not
159 have circulating angiogenic cell measurements, 9 individuals did not provide urine
160 samples, and 8 were missing other covariates. Of the 603 participants with both urinary
161 VOC metabolites and CAC measurements, 228 were excluded because they were
162 current tobacco users. Tobacco use was confirmed using urinary cotinine values, with a
163 value of >40 μg/g creatinine indicating current user. The final sample size consisted of
164 375 participants.

165 **2.2 Urinary Volatile Organic Compound Metabolites**

166 Urinary metabolites of 12 VOCs (16 metabolites), and free forms of the nicotine
167 metabolite cotinine were quantified by ultra-performance liquid chromatography-mass
168 spectrometry (UPLC-MS/MS) as described before (Keith et al., 2020; Lorkiewicz et al.,
169 2019; Yeager et al., 2020). All VOC metabolites that were below the limit of detection
170 (LOD) were imputed with the LOD value divided by the square root of 2. All urinary
171 analytes were normalized to urinary creatinine levels, measured on a COBAS MIRA-plus
172 analyzer (Roche, New Jersey) with Infinity Creatinine Reagent (Thermo Fisher Scientific,
173 Massachusetts). A complete list of metabolites and parent compounds is provided in
174 supplementary material (Table S1).

175 **2.3 Circulating Angiogenic Cells**

176 Fifteen different subtypes of circulating angiogenic cells were measured in the
177 peripheral blood by flow cytometry - on fresh blood specimens as previously described
178 (DeJarnett et al., 2014; O'Toole et al., 2010). CACs were defined as CD14 negative cells
179 and differentiated into subtypes based on expression of CD45, CD34, CD146 and AC133
180 surface markers.

181 **2.4 Air Pollution and Meteorological Data**

182 Data on ambient levels of $PM_{2.5}$ (particulate matter with an aerodynamic diameter
183 of ≤2.5 μm; μg/m^3) and ozone (ppm) were obtained from three regional EPA-validated
184 monitoring stations in the Louisville, KY region. We used the average of the three monitors
185 for the day the study participants provided blood and urine samples.

186 Data on average daily temperature (°F) and relative humidity (%) were obtained
187 from the National Oceanic and Atmospheric Administration's (NOAA) National Centers
188 for Environmental Information. Data was download for the Louisville International Airport
189 monitoring station, which is in close proximity to the study participants' residence, for the
190 day the study participants provided blood and urine sample. This data is publicly available
191 (https://www.ncdc.noaa.gov/cdo-web/datatools/selectlocation).

192 **2.5 Statistical Analysis**

193 Participant characteristics are expressed as mean ± standard deviation (SD) for
194 continuous variables, and frequency (%) for categorical variables. Descriptive statistics
195 were conducted for all VOC metabolites. Pearson correlations were performed across
196 VOC metabolites and daily $PM_{2.5}$, ozone, temperature, and relative humidity
197 (Supplemental Figure 1). Because CACs were positive, and heavily right-skewed,
198 generalized linear models with the gamma distribution and log-link were used to examine
199 the association between CACs and VOC metabolites (Figure 1). Models used natural log-
200 transformed VOC metabolites, and were adjusted based on prior literature showing an
201 association with VOC exposure or CAC levels (Fadini et al., 2008; Fadini et al., 2007;
202 Giannotti et al., 2010; O'Toole et al., 2010; Samman Tahhan et al., 2018; Tobler et al.,

203 2010; Xia et al., 2012). Model adjustments included: age, sex, race, BMI, diabetes status,
204 hypertension status, and daily $PM_{2.5}$ levels. There are a total of 240 unique comparisons
205 between CACs and VOC metabolites. To account for multiple testing, we applied the
206 Benjamini-Hochberg procedure to control the false-discovery rate (FDR). We examined
207 potential non-linear relationships between VOC metabolites and CACs using restricted
208 cubic splines, with tests for non-linearity using the likelihood ratio test. An environmental
209 risk score (ERS) representing cumulative VOC exposure risk was calculated for each
210 CAC using a modified version of the method developed by Park *et al* (2014). Briefly, (1)
211 VOC metabolites were log transformed, (2) each log-transformed VOC metabolite was
212 regressed against each CAC separately, while adjusting for covariates, (3) the
213 environmental risk score for each participant was calculated as the sum of log-
214 transformed VOC metabolite x (regression coefficient/standard error). If multiple VOC
215 metabolites came from the same parent compound, only one was used in the creation of
216 the ERS. The final set of VOC metabolites used to calculate the ERS consist of: 3HPMA,
217 AAMA, CYMA, MU, MHBMA3, HPMMA, AMCC, PGA, 2HPMA, MA, BMA, and
218 3MHA+4MHA. Multi-pollutant models including all VOC metabolites that were
219 independently associated with CACs were further developed to test whether single VOC
220 effects were confounded by correlated VOCs (Pearson correlations provided in
221 Supplemental Table 2). These models used log-transformed metabolite values, and the
222 same adjustments as single pollutant models. Interactions between covariates and log
223 VOC metabolites were tested in single pollutant models. We stratified analyses by
224 participant characteristics to examine whether effect estimates differed by subgroups.

In addition to single-VOC models, we created models for multiple VOC exposures. To assess multiple VOC exposures as a mixture, we applied a Bayesian Kernel Machine Regression (BKMR) approach to flexibly model the relationship between the mixture of VOC metabolites and each of the 15 CACs (Bobb et al., 2015). The BKMR is a semi-parametric technique that models the exposure response relationship as a non-parametric kernel function of the mixture components, adjusting for covariates parametrically. Implementation of BKMR relied on a Gaussian kernel distribution, with flat priors. The Gaussian kernel flexibly captures underlying functional forms without specifying the shape of the exposure-response curve. This method can assess independent effects, estimate the overall mixture, and assess for interactions among mixture components. We used R software (version 4.0.2) and the bkmr package (Bobb et al., 2018) to create posterior inclusion probabilities (PIPs) that quantify the relative importance of each exposure in the model, assess univariate and bivariate associations between mixture components and CACs, assess interactions among VOC metabolites, and estimate the effect of the overall mixture on individual CACs.

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

3. RESULTS

247 The demographic characteristics of study participants are summarized in Table 1.
248 Within the 375 participants of the study, the average age was 50.6 years. A majority were
249 female (64%) and White (83%); 42% were hypertensive, and 17% diabetic. The average
250 daily $PM_{2.5}$ level (±SD) on the day of the study visit was 8.6 ± 3.4 μg/m^3. The urinary
251 metabolites of VOCs measured in study participants are listed in Table 2, along with their
252 parent VOCs.

253 The mean and median concentrations of VOCs show a high degree of variability.
254 Although significant levels of most VOC metabolites could be detected, in many (>20%)
255 samples the level of acrylonitrile (CYMA), styrene/ethylbenzene (PHEMA) metabolites
256 were below the limit of detection (LOD). There was minimal correlation between VOC
257 metabolites and average daily ambient $PM_{2.5}$, ozone, temperature, and relative humidity
258 (Supplemental Figure 1). The levels of urinary VOC metabolites showed a strong
259 dependence on demographics. A comparison of urinary metabolites showed that females
260 had higher levels of the metabolites of N,N'-DMF (AMCC), ethylbenzene/styrene
261 (phenylglyoxylic acid, PGA and mandelic acid, MA) and toluene (BMA) than males.
262 Likewise, in comparison with Blacks, Whites had higher levels of metabolites of benzene
263 (*trans, trans*-muconic acid, MU), N,N'-DMF (N-Acetyl-S-(N-methylcarbamoyl)-L-cysteine,
264 AMCC) and styrene/ethylbenzene (MA). The metabolites of several VOCs – 1,3,
265 butadiene (N-Acetyl-S-(3,4-dihydroxybutyl)-L-cysteine, DHBMA), crotonaldehyde (N-Acetyl-
266 S-(3-hydroxypropyl-1-methyl)-L-cysteine, HPMMA), N,N'-DMF (AMCC),
267 styrene/ethylbenzene (MA), toluene (BMA), and xylene (3MHA+4MHA) were also higher
268 in those >55 years of age. Urinary levels of the metabolite of butadiene (MHBMA3) and
269 toluene (N-Acetyl-S-(benzyl)-L-cysteine, BMA and, 2-Methylhippuric acid, 2MHA) levels were

270 higher in those with a BMI of <30 (Supplemental Table 2), reflecting different levels of
271 exposure.

272 In single-VOC analysis, the levels of several urinary VOC metabolites were
273 associated with the levels of CACs measured in peripheral blood (Figure 1). Some of the
274 most pronounced effects were observed with PGA – a metabolite that could be derived
275 from either ethylbenzene or styrene - which showed a significant inverse association with
276 CACs-1,2,3,4,5,6, and 7, as well as CAC-13 and CAC-15. Several other VOC metabolites
277 showed similar, but variable associations. The percent difference ranged from -10.7%
278 (95% CI: -19.6%, -0.8%) to -42.8% (95% CI: -53.1%, -30.1%). Many VOC metabolites
279 were inversely associated with CAC levels, including metabolites of xylene (3-MHA+4-
280 MHA), acrylonitrile (CYMA), and 1,3-butadiene (MHBMA3). For benzene (MU), we
281 observed a significant inverse association with CAC-1 (-11.9%; 95% CI: -20.0%, -3.0%),
282 and its subpopulation CAC-7 (-10.7%; 95% CI: -18.3%, -2.3%), along with a positive
283 association with CAC-10. Positive associations were also observed with toluene (BMA)
284 and CAC-5, 11, 13, and 14, which share AC133^{+} marker. After correcting for multiple
285 testing, PGA showed statistically significant associations with CAC-1, 3, 5, 6, and 15
286 (Figure 1C). Further, MU remained significantly associated with CAC-10 and BMA
287 remained significantly associated with CAC-5. In our sensitivity analysis, additionally
288 adjusting temperature, relative humidity, and ambient ozone did not substantially alter our
289 associations (Supplemental Figure 2). Overall, these observations suggest that exposure
290 to several VOCs is generally associated with lower levels of several CAC subpopulations
291 in peripheral blood, particularly the early progenitor CACs.

292 To determine the extent to which total VOC exposure can affect CAC levels, and
293 which cell populations contribute to the directionality of the associations (found in Figure
294 1) we created an Environmental Risk Score (ERS) for each cell population (Figure 2). We
295 found a strong negative risk score for CAC-1 (-62.9), CAC-3 (-54.6), and CAC-7 (-62.1),
296 which are a related group of cells with common markers $CD45^{dim}/CD146^{+}/CD34^{+}$. To a
297 lesser extent, we also found negative risk scores for CAC-6 (-47.6), and CAC-9 (-37.7),
298 which are $CD34^{+}$ cells. In comparison, we found positive risk scores for $CD45^{+}$ cells (CAC-
299 12), along with subgroups of $CD45^{+}$ cells: CAC-2, CAC-8, and CAC-14. In addition, we
300 also found a positive score for CAC-10, which is $CD146^{+}$. This indicates that the levels of
301 $CD45^{+}$ cells were positively associated with the urinary metabolites of VOC metabolites,
302 whereas $CD45^{dim}$ and $CD34^{+}$ cells were negatively associated with the metabolite levels.

303 We used restricted cubic splines to determine whether the relationship between
304 VOC metabolites and CACs was non-linear (Figure 3). We observed non-linear
305 relationships between the circulating levels of several cell populations and the metabolite
306 of benzene (MU). In general, low levels of MU was associated with higher levels of
307 circulating cells, whereas higher MU levels were associated with lower cell levels. These
308 non-linear relationships are presented in Figure 3A, which shows a near significant
309 ($p<0.1$) non-linear relationships with CD146+ cell populations (CAC-2, 4, 5, and 6), and
310 with AC133+ cell population – CAC-4 and CAC-5.

311 To determine whether $AC133^{+}$ cells contribute to the non-linear relationships, we
312 examined the association of subgroups of $CD146^{+}$ cells, CAC-4 ($CD45^{+}/CD146^{+}/AC133^{+}$)
313 and CAC-8 ($CD45^{+}/CD146^{+}/AC133^{-}$) with VOC metabolites. We found near significant
314 ($p<0.1$) non-linear relationships between CAC-4 and metabolites of acrolein, 1,3-

315 butadiene, N, N-dimethylformamide, propylene oxide, ethylbenzene/styrene, and xylene.
316 These non-linear relationships were not observed with CAC-8, with the exception of
317 xylene (3-MHA+4-MHA), suggesting that early progenitor endothelial cells
318 ($CD146^+/AC133^+$) are recruited at low exposure levels, but are depleted at high VOC
319 exposures.

320 Given that co-exposure to VOCs is common, we used multi-VOC models to assess
321 whether the significant VOC effects found in single-VOC models were confounded by co-
322 exposures to other VOCs (Figure 4). For this analysis, we focused on rare cell populations
323 with negative ERS scores CAC-3, CAC-5, CAC-7, and CAC-13 (Figure 2). We observed
324 that the metabolite of ethylbenzene/styrene (PGA) remained negatively associated with
325 CAC-3 (-37.0% per 2-fold increase; 95% CI: -48.7%, -22.6%), CAC-5 (-29.8%; 95% CI: -
326 40.6%, -17.0%), CAC-7 (-16.8%; 95% CI: -29.1%, -2.4%), and CAC-13 (-13.1%; 95% CI:
327 -22.2%, -2.8%), suggesting that the associations between PGA and CACs are
328 independent of other VOC metabolite levels. We also found that the metabolites of xylene
329 (2MHA) remained negatively associated with CAC-3, crotonaldehyde (HPMMA)
330 remained positively associated with CAC-5, benzene (MU) remained negatively
331 associated with CAC-7, and toluene (BMA) remained positively associated with CAC-13.

332 In our Bayesian kernel machine regression analysis, important VOC metabolites
333 contributing to changes in CACs are indicated by high posterior inclusion probabilities
334 (Table 3). The directionalities of single VOC associations in our mixture models were
335 generally consistent with the observed directionalities in our single VOC generalized
336 linear models. For example, in mixture models, individuals with greater PGA levels had
337 lower numbers of CACs, while individuals with greater BMA and HPMMA levels had

338 higher numbers of CACs. As evident in Figure 5, we also found that the mixture of VOCs
339 had a nonlinear relationship with CAC-4, similar to the nonlinear associations observed
340 in our single VOC models. We also found that greater exposure to the VOC mixture was
341 associated with higher CAC-10 levels (Figure 5). The effect of the VOC mixture on CAC-
342 10 was largely driven by benzene (MU), with a posterior inclusion probability of 0.55,
343 where all other VOC metabolites were less than 0.15. We found minimal evidence of
344 interactions between pairs of VOC metabolites in our mixture analysis (data not shown).

345 Because the metabolites of benzene, 1,3-butadiene, N,N-dimethylformamide, and
346 ethylbenzene/styrene consistently showed strong associations with CACs, we
347 investigated whether these associations were modified by the demographic
348 characteristics of the participants or their co-morbidity profile. In our subgroup analysis,
349 we found a common theme in many of the models, in which participants who were male,
350 White, normotensive, and recruited on low $PM_{2.5}$ days showed stronger inverse
351 relationships between VOCs and CAC levels (Figure 6).

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

360 We also found some evidence that age and BMI modify the relationships between
361 VOCs and CAC levels, although the direction was not consistent in all models.
362 Interestingly, in several of our models, we found stronger inverse associations between
363 VOC metabolite and CACs in normotensives than hypertensives, and to a lesser extent,
364 non-diabetics than diabetics, - conditions that are known to affect cell migration. For
365 example, for CAC-5, we found a negative association with ethylbenzene/styrene
366 metabolite (PGA) in non-diabetics of -28.2% per 2-fold increase, and with N,N-
367 dimethylformamide (AMCC) in non-diabetics of -24.0% per 2-fold increase, while the
368 relationships in diabetics were not significant. To determine whether these results were
369 due to differential levels of VOC exposure between subgroups, we calculated average
370 metabolite levels for each subgroup *Supplemental Table 2*. We found that levels of MU
371 (benzene) and PGA (ethylbenzene and styrene) were higher in diabetics than non-
372 diabetes and were higher in hypertensive than normotensive individuals, although
373 normotensive had higher levels of 2HPMA (propylene oxide).

374 **4. DISCUSSION**

375 The major findings of this study are that the urinary metabolites of several common
376 VOCs, such as ethylbenzene/styrene, benzene, xylene, and 1,3-butadiene were
377 negatively associated with circulating angiogenic cells. For some VOCs, such as
378 benzene, we observed that low levels of their metabolites were positively associated with
379 an increase in CAC levels, whereas at high levels of exposure were associated with a
380 significant decrease in CAC levels. Our environmental risk score identified that CAC-1,
381 CAC-3, and CAC-7, with common cell markers $CD45^{dim}/CD146^{+}/CD34^{+}$, were the most
382 sensitive to total VOC exposure. Our mixture analysis revealed that the total VOC

383 metabolite mixture was positively associated with CD146^{+} cells. We also found sex, race,
384 age, and hypertension to be effect modifiers of the some of the relationships between the
385 levels of VOC metabolites and CACs. Taken together, these findings suggest that
386 exposure to VOCs could lead to deficits in endothelial repair capacity in a dose-dependent
387 manner, which vary with sex, race, and comorbidities. Given that low CAC levels are
388 predictive of future cardiovascular events and death (Werner et al., 2005), these results
389 support the notion that VOC exposure is associated with significant vascular injury and
390 therefore it may be an important environmental risk factor for CVD.

391 The potential cardiovascular effects of VOCs derived from tobacco smoke have
392 been extensively studied (De Jesús et al., 2020; Keith et al., 2020; Lorkiewicz et al.,
393 2019); however, in contrast, the cardiovascular effects of VOCs not derived from tobacco
394 smoke have received scant attention. Nonetheless, in addition to tobacco smoke,
395 automobile emissions, industrial facilities, hazardous waste sites, and household
396 products are major sources of VOC emissions, and therefore, VOC exposures are
397 common and frequent (Li et al., 2021). In our study, we found significant levels of the
398 VOC metabolites in the urine of non-smokers, indicating particularly high levels of
399 exposure to acrolein, crotonaldehyde, and 1,3-butadiene, and acrylamide, as well as the
400 BTEXS (benzene, toluene, ethylbenzene and styrene) group of VOCs, suggesting a
401 contribution of both indoor and outdoor exposure. Specifically, we found that the CAC
402 levels were particularly sensitive to metabolites of BTEXS compounds. We observed
403 significant associations between CACs and urinary metabolite of benzene (3 CACs),
404 toluene (4 CACs), ethylbenzene/styrene (9 CACs), and xylene (4 CACs).

405 Although exposure to BTEXS is common, little is known about their cardiovascular
406 effects. A study from Hong Kong found that short term elevations in atmospheric levels
407 of benzene and TEX (toluene, ethylbenzene, and xylene) was associated with increases
408 in circulatory disease mortality by 5.8% and 3.5% (Ran et al., 2018a), although in a study
409 from Toronto, no intra-urban variations in VOC levels were found be associated with
410 cardiovascular disease mortality (Villeneuve et al., 2013). Other BTEXS such as toluene,
411 styrene, ethylbenzene, and xylenes have been also linked to a higher prevalence of CVD
412 (Xu et al., 2009), and exposure to VOCs such as benzene, ethylbenzene, and styrene
413 have also been found to be associated with higher odds of developing metabolic
414 syndrome (Shim et al., 2019). Along with these data, the results of our study, support the
415 notion that exposure to VOC could have adverse cardiovascular consequences,
416 potentially due to low-level vascular injury reflected by changes in CAC levels.

417 We found inverse associations between metabolites of acrylonitrile and 1, 3-
418 butadiene and CACs. Although the cardiovascular effects of acrylonitrile are less well
419 studied, exposure to 1,3-butadiene has been found to accelerate the formation of
420 atherosclerotic lesions in animal models, and polymorphisms in phase II enzymes that
421 metabolize 1,3-butadiene are associated with the development of atherosclerosis in
422 humans (Penn and Snyder, 2007). Hence, our observation that urinary metabolites of 1,3-
423 butadiene are associated with CAC levels suggest that the cardiovascular risk of
424 exposure to this VOC may be in part due to depletion of early, progenitor cells, which aid
425 in vascular repair and thereby can prevent the progression of atherosclerotic lesions.
426 Therefore, our observations linking exposure to VOCs such as 1,3-butadiene to
427 concurrent changes in CACs, not only provides a new biological basis for the

428 cardiovascular effects of VOCs, independent of concurrent exposure to $PM_{2.5}$, but also
429 provides new insights into their mechanism of toxicity.

430 In real world scenario, the urban air contains many different VOCs and therefore,
431 we conducted multi-VOC models controlling for all significantly associated metabolites
432 from the single VOC models. In particular, we found that the urinary metabolite PGA
433 remained significantly associated with the levels of $CD45^{dim}/CD146^{+}/CD34^{+}$ cells (CAC-
434 3, CAC-7), $AC133^{+}/CD146^{+}$ cells (CAC-5), and $AC133^{+}/CD34^{+}$ cells (CAC-13). PGA is
435 oxidized from mandelic acid (MA) by alcohol dehydrogenase, and both are considered
436 the major urinary metabolites of styrene and ethylbenzene exposure (Capella et al.,
437 2019). Styrene is mainly used in the manufacturing of plastics, synthetic rubbers, and
438 resins, and is the top released carcinogen in the United States (Mori et al., 2000), while
439 ethylbenzene is found naturally in coal tar and petroleum, and used in the manufacturing
440 of styrene (Capella et al., 2019). The primary route of exposure to styrene and
441 ethylbenzene is through inhalation, contributing up to 99% of total exposure, with 1-2%
442 from food consumption (Tang et al., 2000). We found that the ethylbenzene/styrene
443 metabolite - PGA, but not MA was negatively associated with CACs. This could be due
444 to the chemical structure of PGA, which is a ketone, and is therefore, more reactive than
445 the acid metabolite MA. This could also be representative of higher ethylbenzene/styrene
446 exposure, which would affect the ratio of PGA to MA through an increase in alcohol
447 dehydrogenase activity. The results of our multi-VOC analysis indicate that CACs may
448 be particularly sensitive to ethylbenzene/styrene exposures.

449 We found that the negative association between the N,N-DMF metabolite (AMCC)
450 and CACs was significantly stronger in normotensive participants than those with

451 hypertension. Previous work has shown that hypertension is associated with lower CAC
452 levels and that hypertension is a major independent predictor of impaired endothelial
453 progenitor cell migration (Vasa et al., 2001). It has also been demonstrated that
454 hypertension accelerates EPC senescence in both clinical and experimental settings
455 (Imanishi et al., 2005). These findings suggest that because of the inherent susceptibility
456 of CACs to hypertension, CACs in hypertensive participants may be more vulnerable to
457 exposure to VOCs such as N,N-DMF. N,N-DMF is a commonly used solvent with wide
458 industrial applications, and occupational exposure to high levels of N,N-DMF is
459 associated with liver injury (Li and Zeng, 2019); however, our observation that even
460 ambient exposure to N,N-DMF is associated with vascular injury, suggests that there may
461 be a high risk of hypertension and cardiovascular disease even with low-level exposures.

462 The observed association between VOC metabolites and diminished circulating
463 angiogenic cells suggest that individuals exposed to VOCs may have a lower capacity to
464 repair the endothelium. These effects of VOC exposure appear to be similar to those of
465 other environmental pollutants, which are also associated with suppression of CAC levels.
466 Previous work has shown that exposure to $PM_{2.5}$ over a period of 24h to several weeks
467 suppresses EPC levels (Haberzettl et al., 2012; O'Toole et al., 2010), and exposure to
468 pollutants such as particulate nickel (Liberda et al., 2010), acrolein (DeJarnett et al.,
469 2014) and benzene (Abplanalp et al., 2017) have also been shown to reduce CAC levels.
470 These observations suggest that pollutant exposure either depletes CAC levels in
471 peripheral blood or interferes with the release from the bone marrow and the spleen.
472 However, in contrast to such suppression, CAC levels are slightly elevated in individuals
473 breathing coarse PM (2013a) or those living close to a major roadway (DeJarnett et al.,

474 2015). This biphasic response is consistent with the dose-response relationships
475 observed in the current study, which are indicative of increased CAC mobilization (to
476 promote repair) at low levels of exposure, and CAC suppression at higher exposures.

477 Recently, Yeager *et al.*, have reported that living in areas of high vegetation is
478 associated with increased levels of $CD34^{+}/CD45^{+}/AC133^{+}$ CACs, but decreased
479 $CD31^{+}/CD34^{+}/CD45^{dim}/AC133^{+}$ CACs (2018), suggesting greenness may be protective
480 against air pollution related health effects. Thus, it seems that low levels of toxicity could
481 lead to increased recruitment of CAC into the peripheral blood, whereas, high levels of
482 exposure could suppress CAC recruitment and/or depleting CAC by increasing utilization.
483 The directionality of the effects of exposure on CAC levels may also depend upon locus
484 of injury, some exposures may prevent recruitment, while others may induce injury that
485 stimulates recruitment. Regardless, our results with VOCs suggest that exposure to
486 these chemicals suppresses CAC levels in peripheral blood, but further studies are
487 required to determine whether this is due to decreased recruitment or increased
488 utilization.

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

493 In comparison with $CD45^{dim}$ cells, $CD45^{+}$ CACs are more primitive, with the potential to
494 self-renew and differentiate into specialized cells that can control biological activities such
495 as homeostasis and response to inflammation (Chotinantakul and Leeanansaksiri, 2012).
496 Additionally, the angiogenic cells are rare, and can be easily depleted when used in the

497 repair of endothelial injury. Our exposure response curves are consistent with this notion
498 that at low levels of VOC exposure, $CD45^+$ cells have the ability to regenerate to repair
499 endothelial injury, however at high VOC exposure, the injury surpasses the cells capacity
500 to regenerate or the mobilization of these cells from depots such as the spleen or bone
501 marrow is prevented. Consistent with this notion, results from our Bayesian kernel
502 machine regression also found a nonlinear relationship with the mixture of VOCs and
503 CAC-4 ($CD45^+$/$CD146^+$/$AC133^+$). Further, we found that the mixture of VOC metabolites
504 was positively associated with $CD146^+$ cells. These cells are thought to be mature cells
505 which detach from the intimal monolayer after endothelial injury. Studies have shown an
506 increased number of $CD146^+$ cells in CVD, such as unstable angina, acute myocardial
507 infarction, stroke, and diabetes (Boos et al., 2006). Positive associations could also be
508 due to the specific VOC exposure. We found primarily positive relationships between the
509 toluene metabolite (BMA) and CACs. This could be reflective of the low level of toxicity of
510 toluene in comparison to other VOCs.

511 The major strength of this study is the use of individual level data to assess
512 exposure. Although, measurements of the urinary metabolites do not permit source
513 assignment, they do provide an integrated measure to total exposure of an individual
514 whether via inhalation, ingestion, or endogenous production. Hence, in linking exposure
515 to outcome, such measurements provide a unique advantage as exposure-response
516 could be concurrently evaluated in the same individual. Using this approach, we were
517 able to assess association between exposure and a sensitive and robust measure of
518 preclinical cardiovascular injury (CACs) which is predictive of CVD events and mortality.

519 In our analysis, we also adjusted for exposure to $PM_{2.5}$ and were able to apply a
520 composite VOC risk score that could be used in future studies.

521 Our study also has some limitations. Because of the cross-sectional nature of the
522 study, the associations found in this study cannot address causality. Longitudinal studies
523 are needed to determine whether VOC exposures are associated with adverse clinical
524 outcomes, and whether these outcomes are related to changes in CAC levels. Our study
525 also has the potential for type I errors due to testing of multiple CACs and VOC
526 metabolites. However, due to the correlation between VOC metabolites, and limited prior
527 research on VOC exposure and CVD, we did not adjust for multiple comparisons, to
528 preserve our type II error, and identify potential VOC candidates for future studies.
529 Moreover, we used spot urine collection from each individual to estimate VOC exposure.
530 Although it has been previously shown that spot urine biomarkers of VOC exposure in
531 cigarette smokers correlate well with 24 h urine excretion(Sarkar et al., 2013), our study
532 excluded smokers. Following exposure, many VOCs rapidly undergo biotransformation
533 to generate metabolites that can conjugate with glutathione. Other enzymatic reactions
534 remove glutamic acid and glycine to produce a cysteine conjugate, which is N-acetylated
535 and excreted in the urine as mercapturic acid (Li et al., 2021). It has also been shown that
536 smoking one cigarette, which generates high levels of VOCs, is sufficient to immediately
537 affect the levels of CACs (Mobarrez et al., 2014). Therefore, because VOC metabolites
538 and CAC levels were concurrently measured, it is likely that the observed association
539 represent consequences of recent exposure. Additionally, our urinary VOC metabolites
540 were normalized to creatinine to account for dilution differences, which could introduce
541 error that relates to the dependence of creatinine excretions due to muscle mass, physical

542 activity, and kidney disease. Another limitation of our study is that most participants were,
543 by study design, of older age and at increased risk of cardiovascular disease. Thus, our
544 sample could represent a group of susceptible individuals, which may limit generalizability
545 of our findings.

546 In conclusion, we showed that metabolites of several VOCs were negatively
547 associated with circulating levels of $CD45^{dim}/CD146^{+}/CD34^{+}$ cells and the mixture of VOC
548 metabolites was positively associated with $CD146^{+}$ cells in the peripheral blood. These
549 findings suggest that exposure to low ambient levels of VOCs could damage the
550 endothelium, and deplete rare CACs or inhibit their recruitment, and thereby compromise
551 endothelial repair. Given the global burden of CVD, and the ubiquitous distribution of
552 VOCs, it is imperative to further assess their contribution to CVD risk. Because several
553 commonly used products, such as paints, self-care products, and glue, are important
554 source of VOCs, and that there are several sources of indoor emissions (compressed
555 wood, fire-retardants), this study raises the possibility that environmental exposure to
556 VOCs may be an important determinant of CVD risk. Our work also underscores the need
557 for greater awareness of the potential dangers of VOC exposure and heightened vigilance
558 against such exposures. Future longitudinal and mechanistic studies are needed to
559 confirm current findings. If substantiated, an association between VOC exposure and
560 CVD risk could be included in estimates of CVD risk and appropriate limits and
561 recommendations could be developed to reduce both ambient and indoor VOC
562 exposures.

563

564

565 **FIGURE LEGENDS:**

566

567 **Scheme 1.** Recruitment flow chart.

568

Figure 1. Association between the urinary metabolites of VOCs and circulating levels of CACs(A) and heat map of p-values for pair-wise associations between the urinary metabolites of VOCs and circulating levels of CACs (B and C). (A) The X axis represents % difference (± 95% confidence interval) per 2-fold increase in VOC metabolite for each CAC sub-population (CAC 1-15). Significant associations are shown in red. (B) Raw p-values. (C) False Discovery rate adjusted p-values. 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. Generalized linear models (n=375) were adjusted for age, sex, race, BMI, diabetes, hypertension, and daily $PM_{2.5}$ levels.

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.

569

570 **Figure 2**. Association between Environmental Risk Scores and Circulating Angiogenic Cells.
571 Environmental Risk score was calculated for each cell population as follows: 1. The urinary levels
572 of VOC metabolites were log transformed. 2. Each log transformed VOC metabolite level was
573 regressed against the levels of the cells separately, while adjusting for covariates. 3. ERS was
574 calculated as the sum of the log transformed VOC metabolite levels x (regression
575 coefficient/standard error). Mean and standard deviation of the ERS was calculated for each CAC.
576 Values are mean ± SD of environmental risk score for each cell population.
577

578 **Figure 3**. Exposure-response relationships between selected CACs and VOCs. Non-linear
579 relationships between MU metabolites and CACs (**A**); VOCs and CAC-4 (**B**); and VOCs and CAC-
580 8 (**C**) were examined using restricted cubic splines. Tests for non-linearity used the likelihood ratio
581 test. Models are adjusted for age, sex, race, BMI, diabetes, hypertension, beta-blocker use, and
582 daily $PM_{2.5}$ levels. P-value represents likelihood ratio test for non-linearity.
583

584 **Figure 4.** Association between circulating angiogenic cells and multiple VOCs in the same model.
585 Values represent % difference per 2-fold increase in VOC metabolite. VOC metabolites are
586 ordered by strength of their regression coefficients. Only significant metabolites from one-
587 pollutant models were entered in the model. Models adjusted for age, sex, race, BMI, diabetes,
588 hypertension, and daily $PM_{2.5}$ levels.
589

590 **Figure 5**. Overall association of the VOC metabolite mixture (and 95% credible intervals) with
591 selected CACs, when all exposures are set to a given percentile relative to median. Models
592 (n=375) were adjusted for age, sex, race, BMI, diabetes, hypertension, and daily $PM_{2.5}$ levels.
593

594 **Figure 6**. Subgroup analysis of the relationship between different CAC population with the urinary
595 metabolites of benzene (MU), 1,3-butadiene (MHBMA3), ethylbenzene/styrene (PGA), and N,N-
596 Dimethylformamide (AMCC). Generalized linear models were adjusted for covariates - age, sex,
597 race, BMI, diabetes, hypertension, and daily $PM_{2.5}$ levels, unless the covariates were used for
598 stratification. Significant interaction between subgroups in the full model are indicated by an
599 asterisk.
600

601 **Scheme 2**. Gating Strategy for the Human Circulating Angiogenic Cells. Blue indicates negative
602 environmental risk score; orange indicates positive environmental risk score.
603

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

607

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

610

References

Abplanalp, W., et al., 2017. Benzene exposure is associated with cardiovascular disease risk. PLoS One. 12, e0183602.

Abplanalp, W. T., et al., 2019. Benzene Exposure Induces Insulin Resistance in Mice. Toxicol Sci. 167, 426-437.

Atkinson, R., Arey, J., 2003. Atmospheric degradation of volatile organic compounds. Chemical reviews. 103, 4605-4638.

Awe, S. O., et al., 2006. Acrolein induces vasodilatation of rodent mesenteric bed via an EDHF-dependent mechanism. Toxicol Appl Pharmacol. 217, 266-76.

Bobb, J. F., et al., 2018. Statistical software for analyzing the health effects of multiple concurrent exposures via Bayesian kernel machine regression. Environmental Health. 17, 1-10.

Bobb, J. F., et al., 2015. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. Biostatistics. 16, 493-508.

Boos, C. J., et al., 2006. Circulating Endothelial Cells in Cardiovascular Disease. Journal of the American College of Cardiology. 48, 1538-1547.

Brook, R. D., et al., 2013a. The effect of acute exposure to coarse particulate matter air pollution in a rural location on circulating endothelial progenitor cells: results from a randomized controlled study. Inhalation toxicology. 25, 587-592.

Brook, R. D., et al., 2013b. Long-term fine particulate matter exposure and mortality from diabetes mellitus in Canada. Diabetes care. DC_122189.

Brook, R. D., et al., 2010. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation. 121, 2331-2378.

Capella, K. M., et al., 2019. Ethylbenzene and styrene exposure in the United States based on urinary mandelic acid and phenylglyoxylic acid: NHANES 2005–2006 and 2011–2012. Environmental Research. 171, 101-110.

Chotinantakul, K., Leeanansaksiri, W., 2012. Hematopoietic Stem Cell Development, Niches, and Signaling Pathways. Bone Marrow Research. 2012, 16.

Cohen, A. J., et al., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. The Lancet. 389, 1907-1918.

Conklin, D. J., et al., 2010. Acrolein consumption induces systemic dyslipidemia and lipoprotein modification. Toxicol Appl Pharmacol. 243, 1-12.

Conklin, D. J., et al., 2017. Biomarkers of Chronic Acrolein Inhalation Exposure in Mice: Implications for Tobacco Product-Induced Toxicity. Toxicol Sci. 158, 263-274.

646 De Jesús, V. R., et al., 2020. Urinary Biomarkers of Exposure to Volatile Organic Compounds
647 from the Population Assessment of Tobacco and Health Study Wave 1 (2013-2014). Int J
648 Environ Res Public Health. 17.

649 DeJarnett, N., et al., 2014. Acrolein exposure is associated with increased cardiovascular disease
650 risk. J Am Heart Assoc. 3.

651 DeJarnett, N., et al., 2015. Residential Proximity to Major Roadways Is Associated With Increased
652 Levels of AC133+ Circulating Angiogenic Cells. Arterioscler Thromb Vasc Biol.

653 Fadini, G. P., et al., 2008. Gender differences in endothelial progenitor cells and cardiovascular
654 risk profile: the role of female estrogens. Arteriosclerosis, thrombosis, and vascular
655 biology. 28, 997-1004.

656 Fadini, G. P., et al., 2007. Significance of Endothelial Progenitor Cells in Subjects With Diabetes.
657 Diabetes Care. 30, 1305-1313.

658 Giannotti, G., et al., 2010. Impaired Endothelial Repair Capacity of Early Endothelial Progenitor
659 Cells in Prehypertension. Hypertension. 55, 1389-1397.

660 Haberzettl, P., et al., 2012. Exposure to ambient air fine particulate matter prevents VEGF-
661 induced mobilization of endothelial progenitor cells from the bone marrow. Environmental
662 health perspectives. 120, 848.

663 Heitzer, T., et al., 1996. Cigarette smoking potentiates endothelial dysfunction of forearm
664 resistance vessels in patients with hypercholesterolemia: role of oxidized LDL. Circulation.
665 93, 1346-1353.

666 Hill, J. M., et al., 2003. Circulating endothelial progenitor cells, vascular function, and
667 cardiovascular risk. New England Journal of Medicine. 348, 593-600.

668 Imanishi, T., et al., 2005. Endothelial progenitor cell senescence is accelerated in both
669 experimental hypertensive rats and patients with essential hypertension. Journal of
670 Hypertension. 23, 1831-1837.

671 Ismahil, M. A., et al., 2011. Chronic oral exposure to the aldehyde pollutant acrolein induces
672 dilated cardiomyopathy. Am J Physiol Heart Circ Physiol. 301, H2050-60.

673 Jin, L., et al., 2019. Formaldehyde Induces Mesenteric Artery Relaxation via a Sensitive Transient
674 Receptor Potential Ankyrin-1 (TRPA1) and Endothelium-Dependent Mechanism: Potential
675 Role in Postprandial Hyperemia. Front Physiol. 10, 277.

676 Keith, R. J., et al., 2020. Characterization of Volatile Organic Compound Metabolites in Cigarette
677 Smokers, Electronic Nicotine Device Users, Dual Users, and Nonusers of Tobacco.
678 Nicotine Tob Res. 22, 264-272.

679 Konkle, S. L., et al., 2020. National secular trends in ambient air volatile organic compound levels
680 and biomarkers of exposure in the United States. Environ Res. 182, 108991.

681 Li, A. J., et al., 2021. A review of environmental occurrence, toxicity, biotransformation and
682 biomonitoring of volatile organic compounds. Environmental Chemistry and
683 Ecotoxicology.

684 Li, M. J., Zeng, T., 2019. The deleterious effects of N,N-dimethylformamide on liver: A mini-review.
685 Chem Biol Interact. 298, 129-136.

686 Liberda, E. N., et al., 2010. Exposure to inhaled nickel nanoparticles causes a reduction in number
687 and function of bone marrow endothelial progenitor cells. Inhalation toxicology. 22, 95-99.

688 Lorkiewicz, P., et al., 2019. Comparison of Urinary Biomarkers of Exposure in Humans Using
689 Electronic Cigarettes, Combustible Cigarettes, and Smokeless Tobacco. Nicotine Tob
690 Res. 21, 1228-1238.

691 Lynch, J., et al., 2020. Acute and chronic vascular effects of inhaled crotonaldehyde in mice: Role
692 of TRPA1. Toxicol Appl Pharmacol. 402, 115120.

693 Matanoski, G. M., Tao, X. G., 2003. Styrene exposure and ischemic heart disease: a case-cohort
694 study. Am J Epidemiol. 158, 988-95.

695 McGraw, K. E., et al., 2021. Exposure to volatile organic compounds – acrolein, 1,3-butadiene,
696 and crotonaldehyde – is associated with vascular dysfunction. Environmental Research.
697 196, 110903.

698 Mobarrez, F., et al., 2014. The effects of smoking on levels of endothelial progenitor cells and
699 microparticles in the blood of healthy volunteers. PloS one. 9, e90314-e90314.

700 Mori, T. A., et al., 2000. Purified eicosapentaenoic and docosahexaenoic acids have differential
701 effects on serum lipids and lipoproteins, LDL particle size, glucose, and insulin in mildly
702 hyperlipidemic men. The American Journal of Clinical Nutrition. 71, 1085-1094.

703 Mustafić, H., et al., 2012. Main air pollutants and myocardial infarction: A systematic review and
704 meta-analysis. JAMA. 307, 713-721.

705 O'Toole, T. E., et al., 2010. Episodic exposure to fine particulate air pollution decreases circulating
706 levels of endothelial progenitor cells. Circ Res. 107, 200-3.

707 Park, S. K., et al., 2014. Environmental risk score as a new tool to examine multi-pollutants in
708 epidemiologic research: an example from the NHANES study using serum lipid levels.
709 PloS one. 9, e98632.

710 Penn, A., Snyder, C. A., 2007. 1,3-Butadiene exposure and cardiovascular disease. Mutat Res.
711 621, 42-9.

712 Pittilo, M., 2000. Cigarette smoking, endothelial injury and cardiovascular disease. International
713 Journal of Experimental Pathology. 81, 219-230.

714 Rajagopalan, S., et al., 2018. Air Pollution and Cardiovascular Disease: JACC State-of-the-Art
715 Review. Journal of the American College of Cardiology. 72, 2054-2070.

716 Rajagopalan, S., Brook, R. D., 2012. Air pollution and type 2 diabetes: mechanistic insights.
717 Diabetes. 61, 3037-3045.

718 Ran, J., et al., 2018a. Short-term effects of ambient benzene and TEX (toluene, ethylbenzene,
719 and xylene combined) on cardiorespiratory mortality in Hong Kong. Environ Int. 117, 91-
720 98.

721 Ran, J., et al., 2018b. Are ambient volatile organic compounds environmental stressors for heart
722 failure? Environmental Pollution. 242, 1810-1816.

723 Samman Tahhan, A., et al., 2018. Circulating Progenitor Cells and Racial Differences. Circulation
724 research. 123, 467-476.

725 Sarkar, M., et al., 2013. Evaluation of spot urine as an alternative to 24 h urine collection for
726 determination of biomarkers of exposure to cigarette smoke in adult smokers.
727 Environmental toxicology and pharmacology. 36, 108-114.

728 Shah, A. S. V., et al., 2015. Short term exposure to air pollution and stroke: systematic review
729 and meta-analysis. BMJ. 350.

730 Shim, Y. H., et al., 2019. Association between Heavy Metals, Bisphenol A, Volatile Organic
731 Compounds and Phthalates and Metabolic Syndrome. International Journal of
732 Environmental Research and Public Health. 16, 671.

733 Shuai, J., et al., 2018. Health risk assessment of volatile organic compounds exposure near
734 Daegu dyeing industrial complex in South Korea. BMC Public Health. 18, 528.

735 Sithu, S. D., et al., 2010. Exposure to acrolein by inhalation causes platelet activation. Toxicol
736 Appl Pharmacol. 248, 100-10.

737 Srivastava, S., et al., 2011. Oral exposure to acrolein exacerbates atherosclerosis in apoE-null
738 mice. Atherosclerosis. 215, 301-8.

739 Tang, W., et al., 2000. Estimation of human exposure to styrene and ethylbenzene. Toxicology.
740 144, 39-50.

741 Tobler, K., et al., 2010. Reduction of both number and proliferative activity of human endothelial
742 progenitor cells in obesity. International Journal of Obesity. 34, 687-700.

743 Törnqvist, H., et al., 2007. Persistent Endothelial Dysfunction in Humans after Diesel Exhaust
744 Inhalation. American Journal of Respiratory and Critical Care Medicine. 176, 395-400.

745 Vasa, M., et al., 2001. Number and migratory activity of circulating endothelial progenitor cells
746 inversely correlate with risk factors for coronary artery disease. Circulation research. 89,
747 e1-e7.

748 Villeneuve, P. J., et al., 2013. A cohort study of intra-urban variations in volatile organic
749 compounds and mortality, Toronto, Canada. Environ Pollut. 183, 30-9.

750 Wang, G. W., et al., 2008. Acrolein consumption exacerbates myocardial ischemic injury and
751 blocks nitric oxide-induced PKCepsilon signaling and cardioprotection. J Mol Cell Cardiol.
752 44, 1016-22.

753 Werner, N., et al., 2005. Circulating endothelial progenitor cells and cardiovascular outcomes.
754 New England Journal of Medicine. 353, 999-1007.

755 Werner, N., Nickenig, G., 2006. Influence of cardiovascular risk factors on endothelial progenitor
756 cells: limitations for therapy? Arteriosclerosis, thrombosis, and vascular biology. 26, 257-
757 266.

758 Xia, W. H., et al., 2012. Age-Related Decline in Reendothelialization Capacity of Human
759 Endothelial Progenitor Cells Is Restored by Shear Stress. Hypertension. 59, 1225-1231.

760 Xu, X., et al., 2009. Association between exposure to alkylbenzenes and cardiovascular disease
761 among National Health and Nutrition Examination Survey (NHANES) participants. Int J
762 Occup Environ Health. 15, 385-91.

763 Yang, B.-Y., et al., 2018. Global association between ambient air pollution and blood pressure: A
764 systematic review and meta-analysis. Environmental Pollution. 235, 576-588.

765 Ye, D., et al., 2017. Estimating Acute Cardiorespiratory Effects of Ambient Volatile Organic
766 Compounds. Epidemiology. 28, 197-206.

767 Yeager, R., et al., 2020. Association between residential greenness and exposure to volatile
768 organic compounds. Sci Total Environ. 707, 135435.

769 Yeager, R., et al., 2018. Association Between Residential Greenness and Cardiovascular Disease
770 Risk. J Am Heart Assoc. 7, e009117.

771 Zafar, N., et al., 2018. Circulating angiogenic stem cells in type 2 diabetes are associated with
772 glycemic control and endothelial dysfunction. PLoS One. 13, e0205851.

773

774

775

776

777

778

779

780

781

782

783

784 Table 1. Participant Characteristics

Variable	Overall (n=375)
Sex, Male	135 (36%)
Race	
White	311 (83%)
Black	43 (11%)
Other	21 (6%)
Age (years)	50.6 ± 13.2
BMI (kg/m^2)	30.7 ± 7.0
SBP (mmHg)	128.5 ± 18.1
DBP (mmHg)	78.1 ± 11.2
Hypertension	156 (42%)
Diabetes	65 (17%)
MI	24 (7%)
Medication Use	
ASA	46 (13%)
Statin	60 (17%)
ACE or ARB	63 (18%)
Beta Blockers	35 (10%)

785

786 Values represent mean ± S.D. for continuous variables, and frequency (%) for categorical
787 variables. Abbreviations: BMI indicates body mass index; SBP, systolic blood pressure; DBP,
788 diastolic blood pressure; MI, myocardial infarction; ASA, aspirin; ACE, angiotensin converting
789 enzyme inhibitor; ARB, angiotensin receptor blocker.

790

791

792

793

794

795

796

797

798

799

800

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

Parent Structure	VOC	Metabolite	Mean ± SD	Median±IQR	Min-Max	n (%) LOD
	Acrolein	CEMA	115.8 ± 117	89.6 ± 87.5	3.8 - 1213.7	42 (11.2%)
		3HPMA	341.5 ± 344.6	231.9 ± 222.5	46.3 - 2908.4	0 (0%)
	Acrylamide	AAMA	72.4 ± 61.6	56.2 ± 54	8.3 - 677.7	1 (0.3%)
	Acrylonitrile	CYMA	5.7 ± 17.2	1.3 ± 2.4	0.2 - 207.2	231 (61.6%)
	Benzene	MU	133.4 ± 214	56.5 ± 117.1	6 - 1827	6 (1.6%)
	1,3-Butadiene	DHBMA	353.6 ± 175.9	318.8 ± 192.4	50.2 - 1384.6	0 (0%)
		MHBMA3	9.1 ± 8.2	7.2 ± 5.3	0.8 - 75.4	18 (4.8%)
	Crotonaldehyde	HPMMA	308.7 ± 306.2	226.8 ± 172.6	13.5 - 3589.6	1 (0.3%)
	N,N-Dimethylformamide	AMCC	160.4 ± 117.3	134.8 ± 123.3	2.6 - 864.4	5 (1.3%)
	Ethylbenzene, styrene	PGA	225.9 ± 145.5	204.7 ± 146.9	9.8 - 1437.3	11 (2.9%)
	Propylene oxide	2HPMA	58.4 ± 122.9	27.2 ± 30.5	3.4 - 1044.7	4 (1.1%)
	Styrene, ethylbenzene	PHEMA	2.2 ± 4.1	1.4 ± 1.6	0.2 - 52.6	174 (46.4%)
		MA	232.8 ± 165.5	199.6 ± 132.2	20.9 - 1593.3	1 (0.3%)
	Toluene	BMA	11.5 ± 12.1	7.7 ± 8.1	0.1 - 106.9	4 (1.1%)
	Xylene	2MHA	32.8 ± 75.4	13 ± 27.3	0.2 - 1058.9	53 (14.1%)
		3MHA+4MHA	209.1 ± 409	95.8 ± 142.6	3.4 - 5499.1	30 (8%)

803

804 VOC metabolites were measured in 375 non-smoking participants, 25-70 years of age with low
805 to severe CVD risk. IQR, interquartile range; LOD, limit of detection. The complete chemical
806 names of each metabolite are listed in Supplemental Table 1.

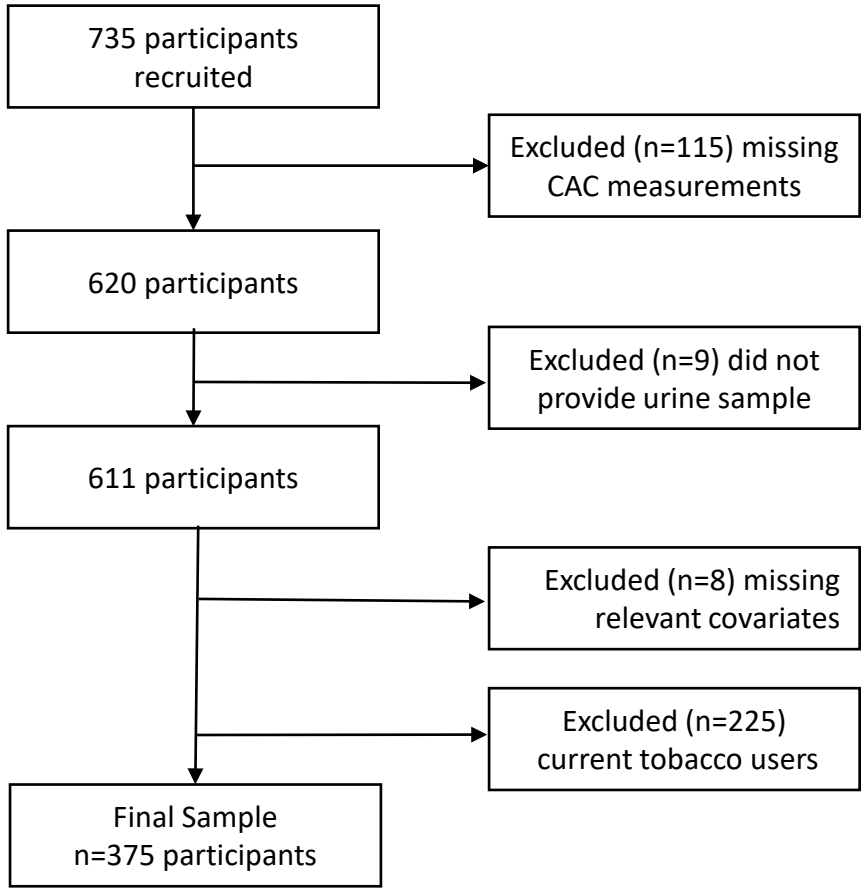
807

808 **Table 3.** Posterior inclusion probabilities of individual VOC metabolites in the BKMR
809 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

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

810


735 participants recruited
Excluded (n=115) missing CAC measurements
620 participants
Excluded (n=9) did not provide urine sample
611 participants
Excluded (n=8) missing relevant covariates
Excluded (n=225) current tobacco users
Final Sample n=375 participants

Scheme 1

Figure 1

A

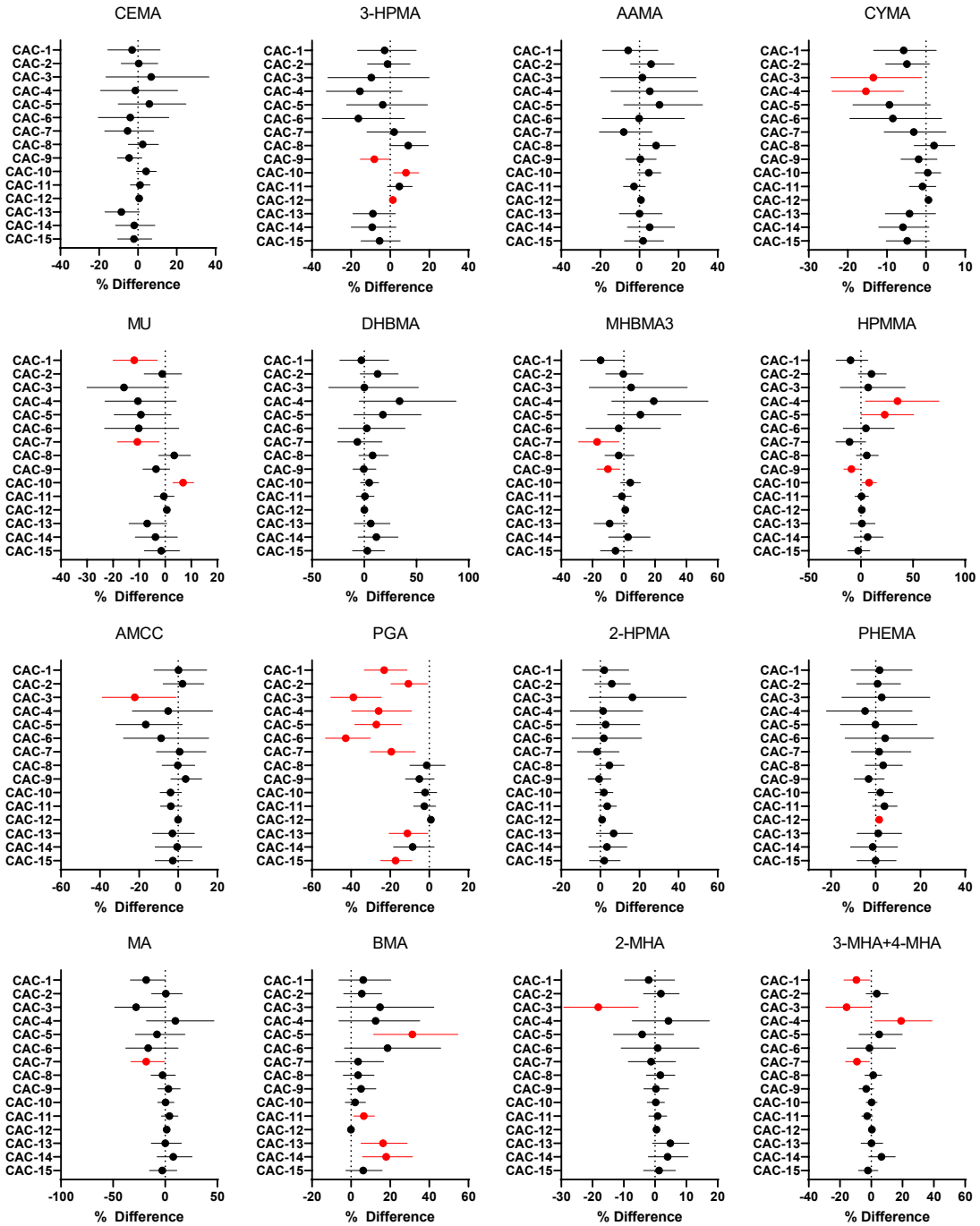
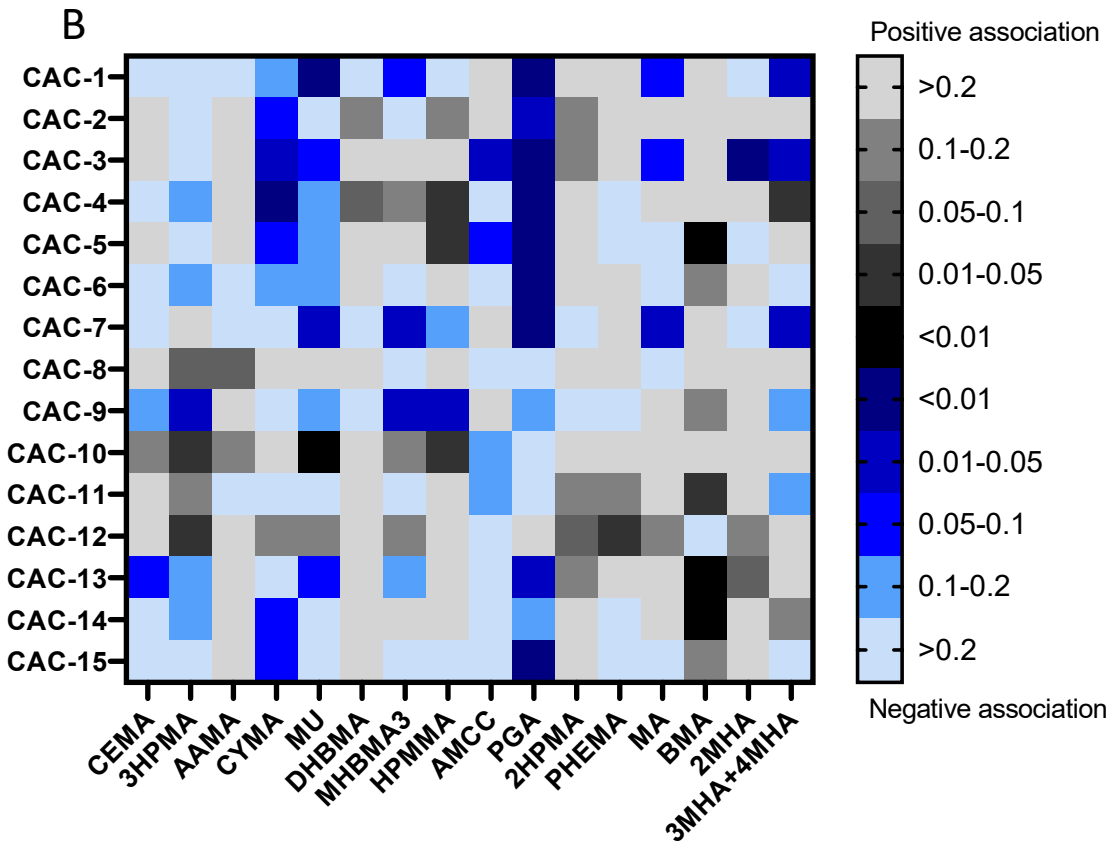

CEMA
3-HPMA
AAMA
CYMA
MU
DHBMA
MHBMA3
HPMMA
AMCC
PGA
2-HPMA
PHEMA
MA
BMA
2-MHA
3-MHA+4-MHA
CAC-1
CAC-2
CAC-3
CAC-4
CAC-5
CAC-6
CAC-7
CAC-8
CAC-9
CAC-10
CAC-11
CAC-12
CAC-13
CAC-14
CAC-15
% Difference

Figure 1


B
CAC-1
CAC-2
CAC-3
CAC-4
CAC-5
CAC-6
CAC-7
CAC-8
CAC-9
CAC-10
CAC-11
CAC-12
CAC-13
CAC-14
CAC-15
CEMA
3HPMA
AAMA
CYMA
MU
DHBMA
MHBMA3
HPMMA
AMCC
PGA
2HPMA
PHEMA
MA
BMA
2MHA
3MHA+4MHA
Positive association
>0.2
0.1-0.2
0.05-0.1
0.01-0.05
<0.01
<0.01
0.01-0.05
0.05-0.1
0.1-0.2
>0.2
Negative association

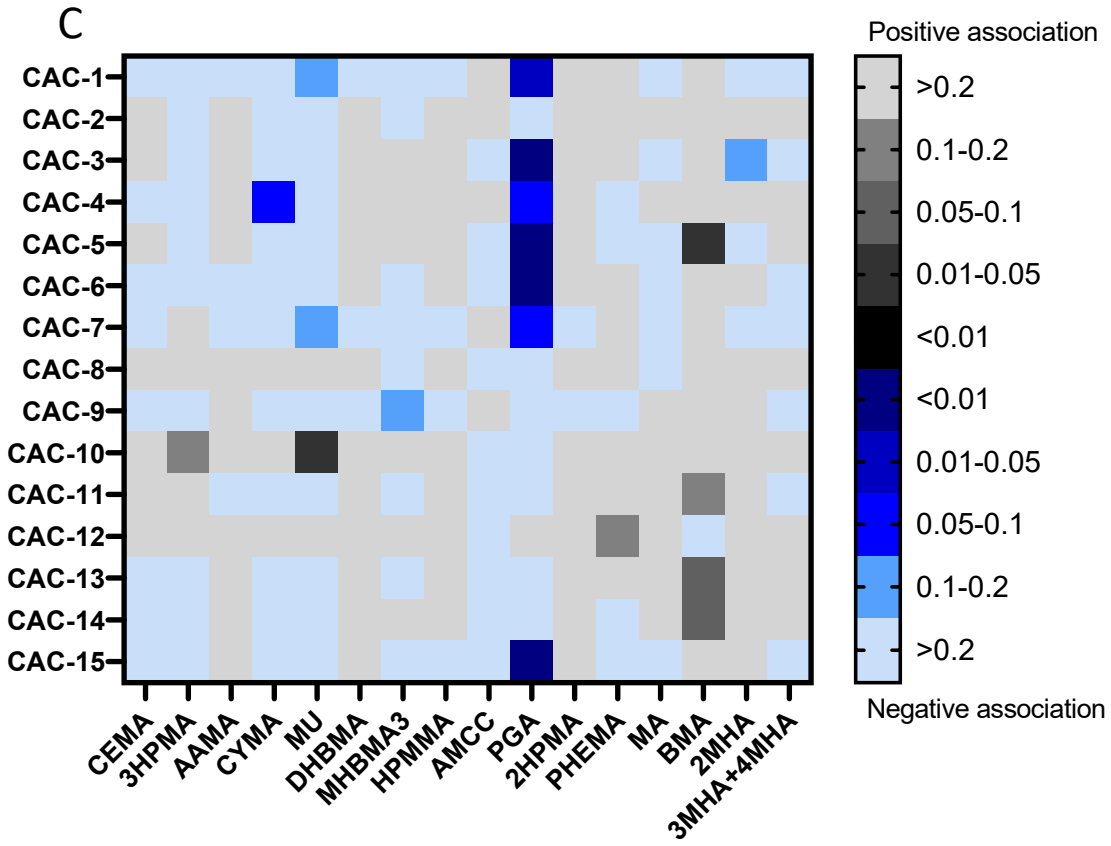

C
CAC-1
CAC-2
CAC-3
CAC-4
CAC-5
CAC-6
CAC-7
CAC-8
CAC-9
CAC-10
CAC-11
CAC-12
CAC-13
CAC-14
CAC-15
CEMA
3HPMA
AAMA
CYMA
MU
DHBMA
MHBMA3
HPMMA
AMCC
PGA
2HPMA
PHEMA
MA
BMA
2MHA
3MHA+4MHA
Positive association
>0.2
0.1-0.2
0.05-0.1
0.01-0.05
<0.01
<0.01
0.01-0.05
0.05-0.1
0.1-0.2
>0.2
Negative association

Figure 2

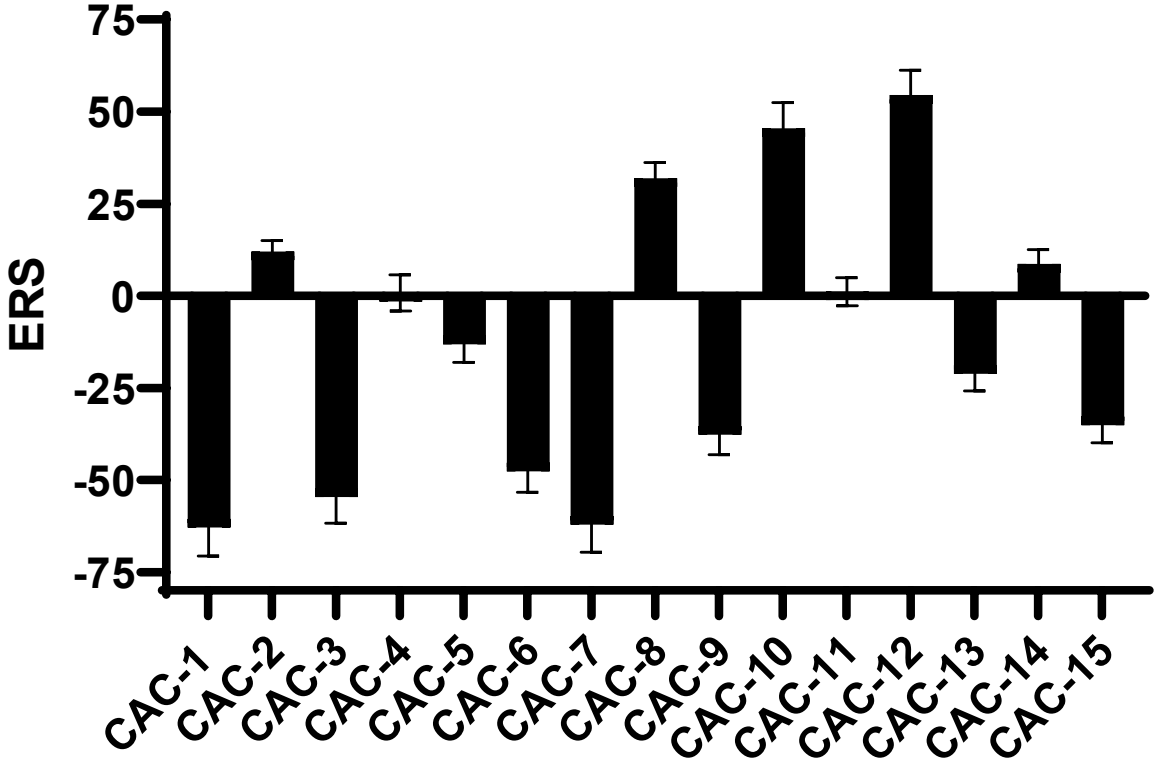

75
50
25
0
-25
-50
-75
ERS
CAC-1
CAC-2
CAC-3
CAC-4
CAC-5
CAC-6
CAC-7
CAC-8
CAC-9
CAC-10
CAC-11
CAC-12
CAC-13
CAC-14
CAC-15

Figure 3

A

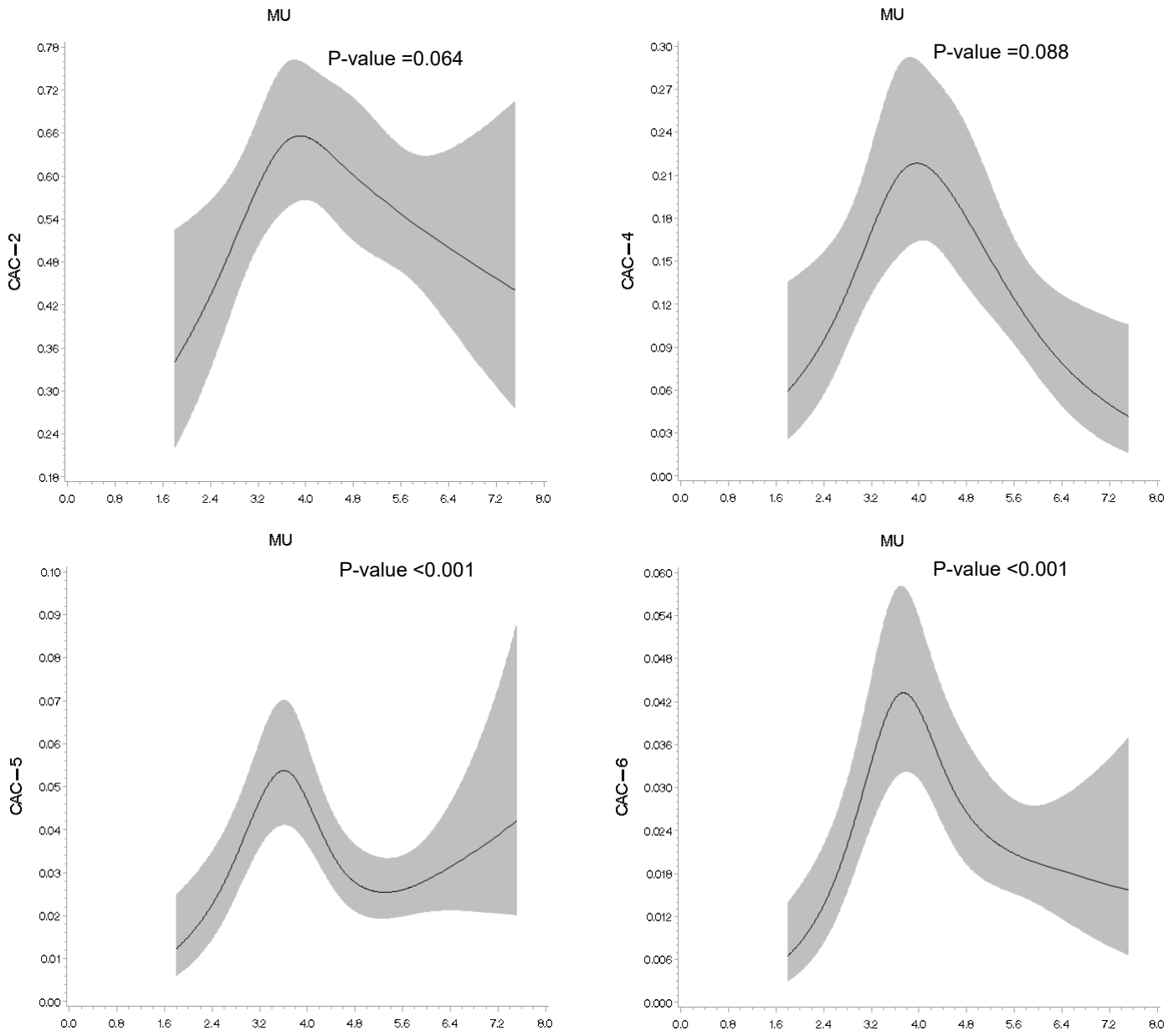

MU
P-value =0.064
CAC−2
0.78
0.72
0.66
0.60
0.54
0.48
0.42
0.36
0.30
0.24
0.18
0.0 0.8 1.6 2.4 3.2 4.0 4.8 5.6 6.4 7.2 8.0
MU
P-value =0.088
CAC−4
0.30
0.27
0.24
0.21
0.18
0.15
0.12
0.09
0.06
0.03
0.00
0.0 0.8 1.6 2.4 3.2 4.0 4.8 5.6 6.4 7.2 8.0
MU
P-value <0.001
CAC−5
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0.00
0.0 0.8 1.6 2.4 3.2 4.0 4.8 5.6 6.4 7.2 8.0
MU
P-value <0.001
CAC−6
0.060
0.054
0.048
0.042
0.036
0.030
0.024
0.018
0.012
0.006
0.000
0.0 0.8 1.6 2.4 3.2 4.0 4.8 5.6 6.4 7.2 8.0

Figure 3

B

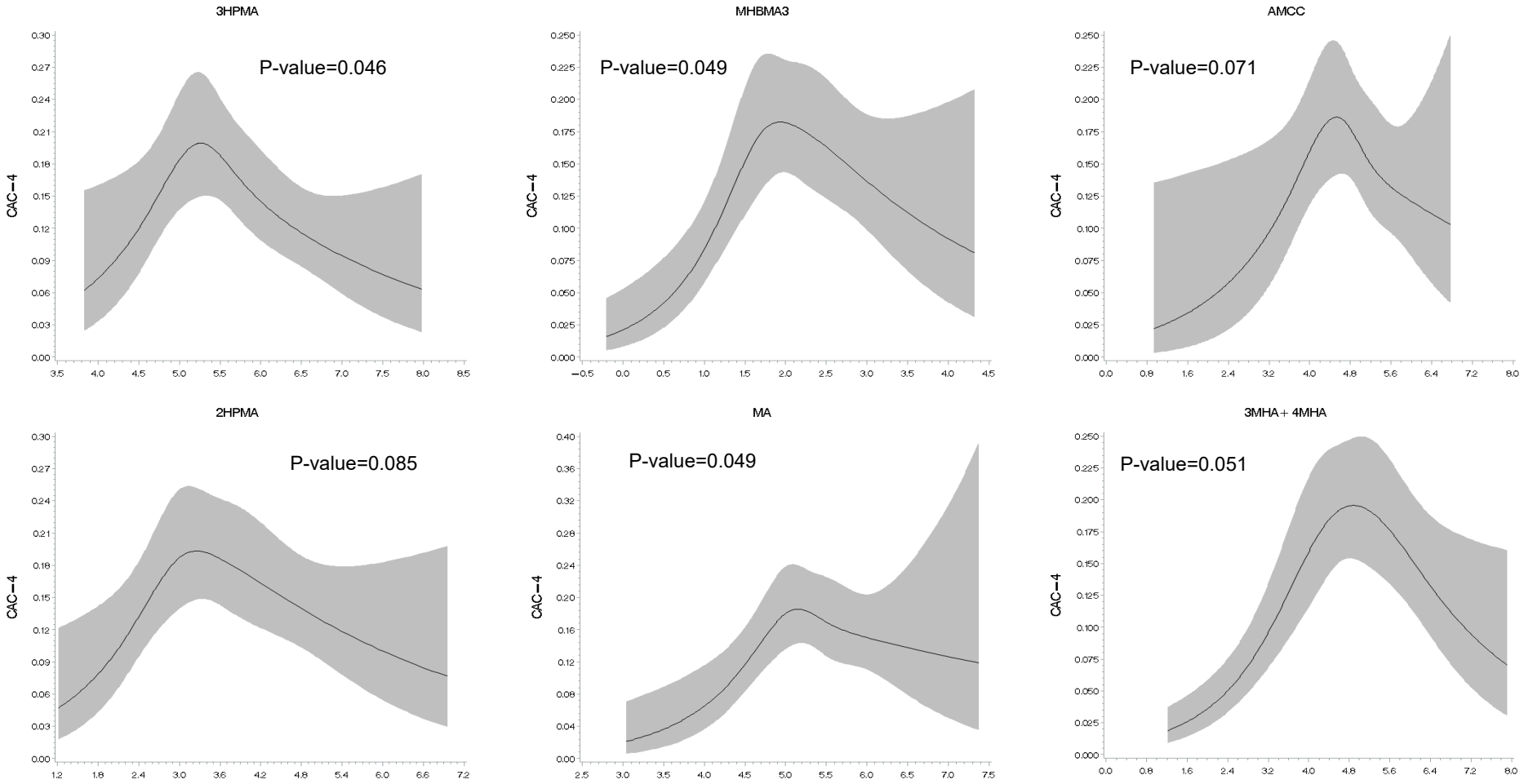

3HPMA
P-value=0.046
MHBMA3
P-value=0.049
AMCC
P-value=0.071
2HPMA
P-value=0.085
MA
P-value=0.049
3MHA+4MHA
P-value=0.051
CAC−4

Figure 3

C

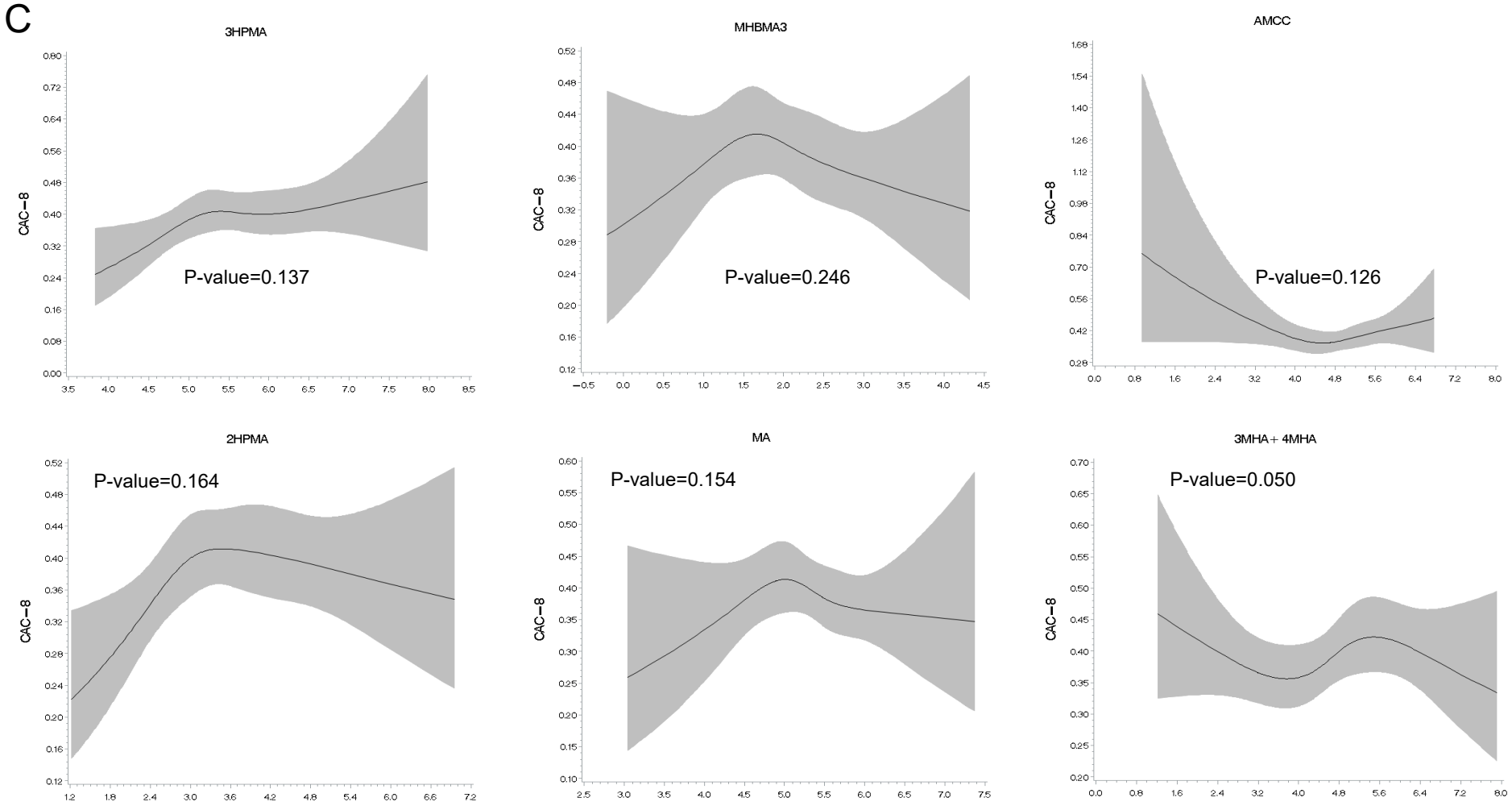

3HPMA
CAC−8
P-value=0.137
MHBMA3
CAC−8
P-value=0.246
AMCC
CAC−8
P-value=0.126
2HPMA
CAC−8
P-value=0.164
MA
CAC−8
P-value=0.154
3MHA+4MHA
CAC−8
P-value=0.050

Figure 4

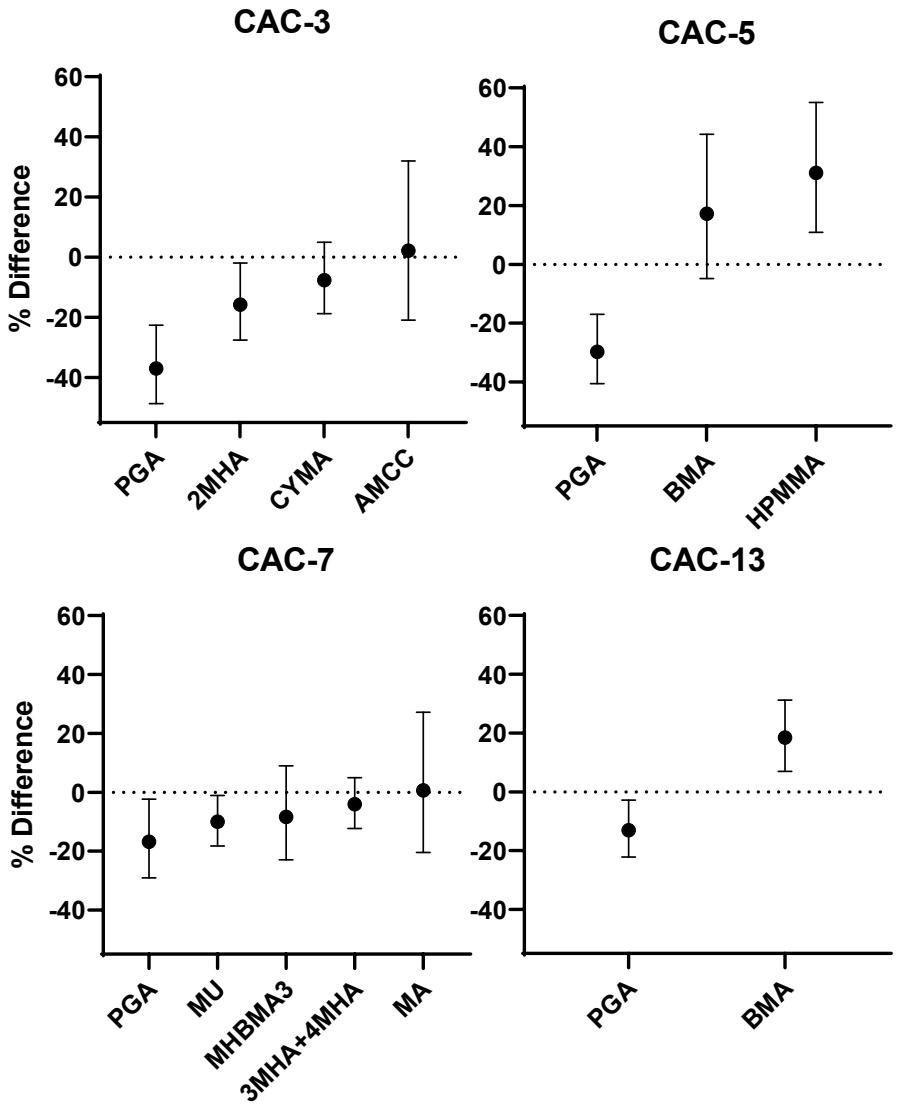

CAC-3
CAC-5
CAC-7
CAC-13
% Difference
60
40
20
0
-20
-40
PGA
2MHA
CYMA
AMCC
BMA
HPMMA
MU
MHBMA3
3MHA+4MHA
MA

Figure 5

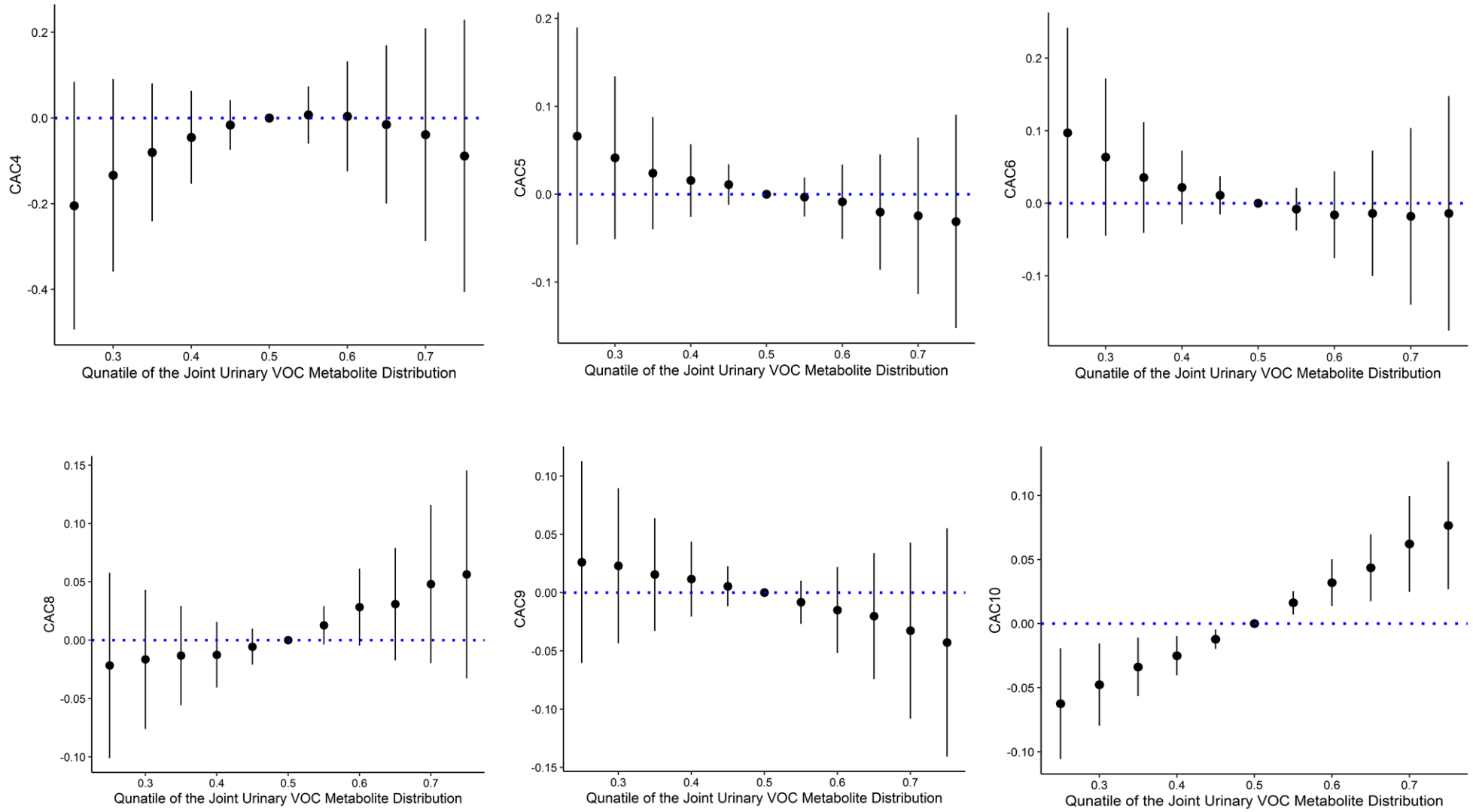
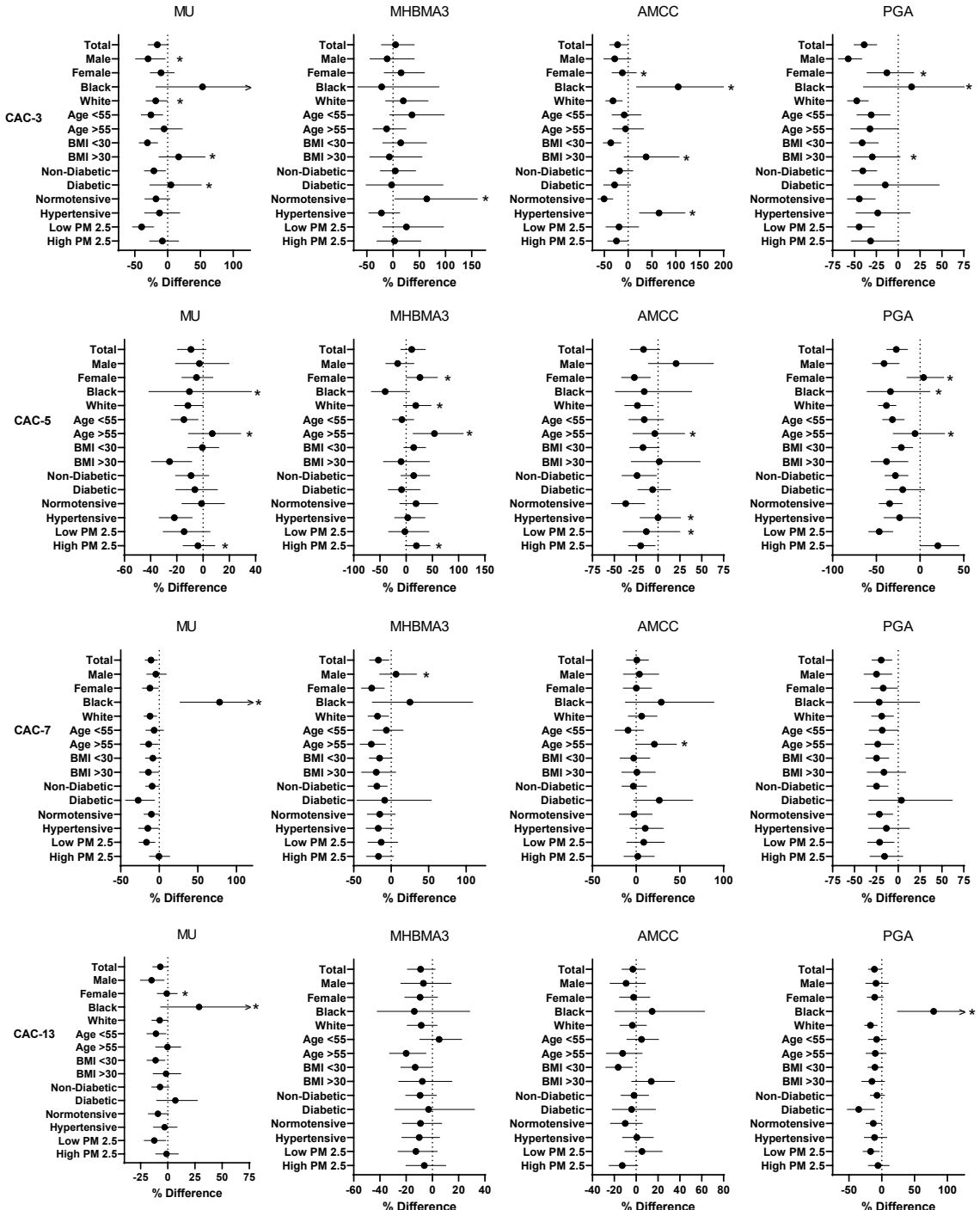
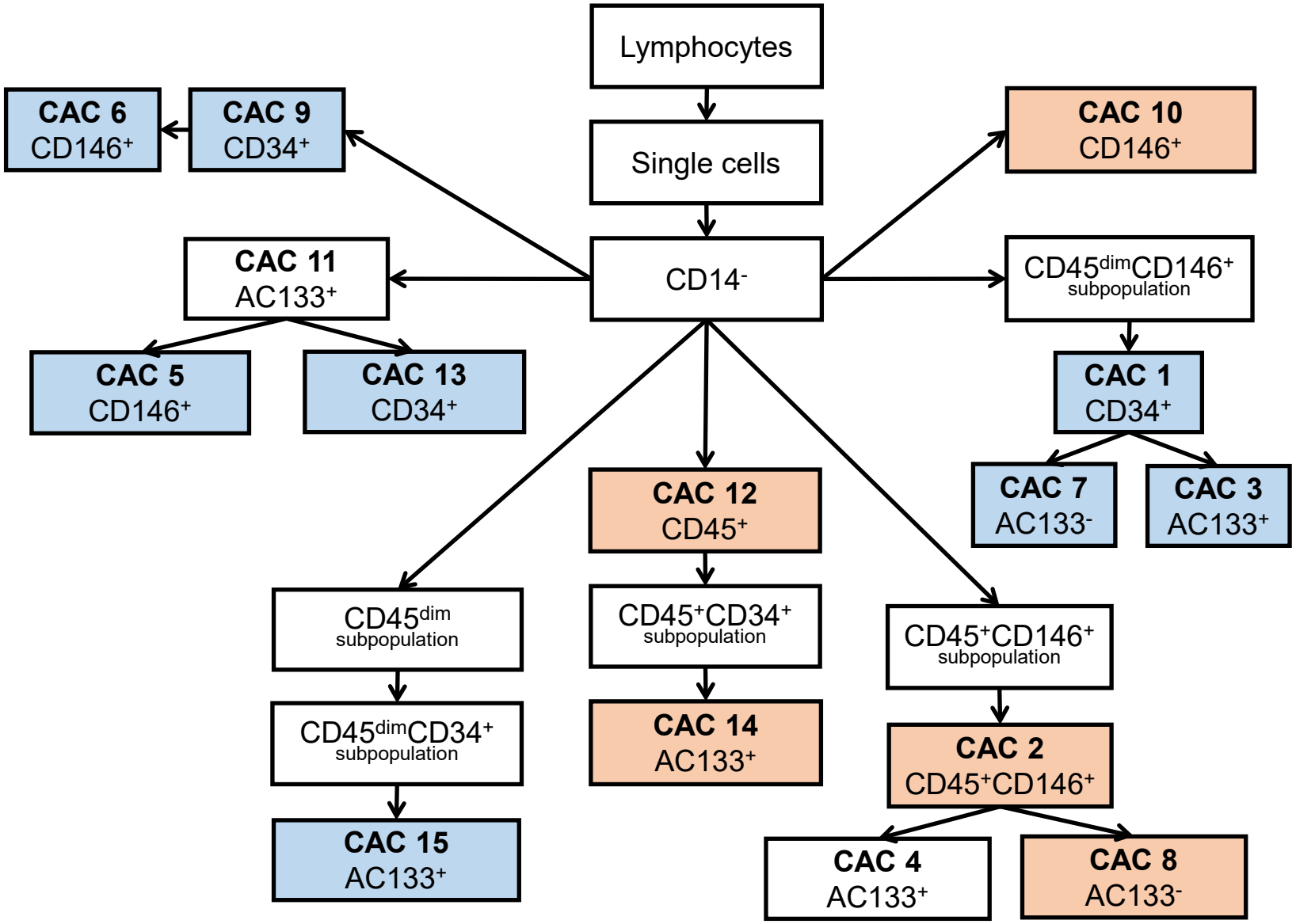

CAC4
0.2
0.0
-0.2
-0.4
0.3
0.4
0.5
0.6
0.7
Qunatile of the Joint Urinary VOC Metabolite Distribution
CAC5
0.2
0.1
0.0
-0.1
0.3
0.4
0.5
0.6
0.7
Qunatile of the Joint Urinary VOC Metabolite Distribution
CAC6
0.2
0.1
0.0
-0.1
0.3
0.4
0.5
0.6
0.7
Qunatile of the Joint Urinary VOC Metabolite Distribution
CAC8
0.15
0.10
0.05
0.00
-0.05
-0.10
0.3
0.4
0.5
0.6
0.7
Qunatile of the Joint Urinary VOC Metabolite Distribution
CAC9
0.10
0.05
0.00
-0.05
-0.10
-0.15
0.3
0.4
0.5
0.6
0.7
Qunatile of the Joint Urinary VOC Metabolite Distribution
CAC10
0.10
0.05
0.00
-0.05
-0.10
0.3
0.4
0.5
0.6
0.7
Qunatile of the Joint Urinary VOC Metabolite Distribution

Figure 6

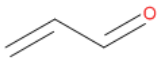
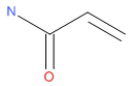
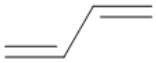
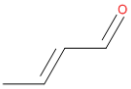
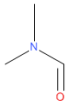
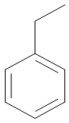
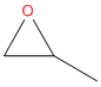
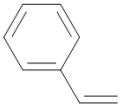
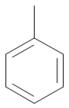
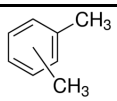

MU
MHBMA3
AMCC
PGA
CAC-3
CAC-5
CAC-7
CAC-13
Total
Male
Female
Black
White
Age <55
Age >55
BMI <30
BMI >30
Non-Diabetic
Diabetic
Normotensive
Hypertensive
Low PM 2.5
High PM 2.5
% Difference


Lymphocytes
Single cells
CD14-
CAC 6
CD146+
CAC 9
CD34+
CAC 10
CD146+
CAC 11
AC133+
CD45dimCD146+
subpopulation
CAC 5
CD146+
CAC 13
CD34+
CAC 1
CD34+
CAC 7
AC133-
CAC 3
AC133+
CAC 12
CD45+
CD45dim
subpopulation
CD45+CD34+
subpopulation
CD45+CD146+
subpopulation
CD45dimCD34+
subpopulation
CAC 14
AC133+
CAC 2
CD45+CD146+
CAC 15
AC133+
CAC 4
AC133+
CAC 8
AC133-

Scheme 2

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-carbamoylethyl)-L-cysteine	AAMA	3.03
	Acrylonitrile	N-Acetyl-S-(2-cyanoethyl)-L-cysteine	CYMA	1.25
	Benzene	*trans, trans*-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 for specific subgroups.

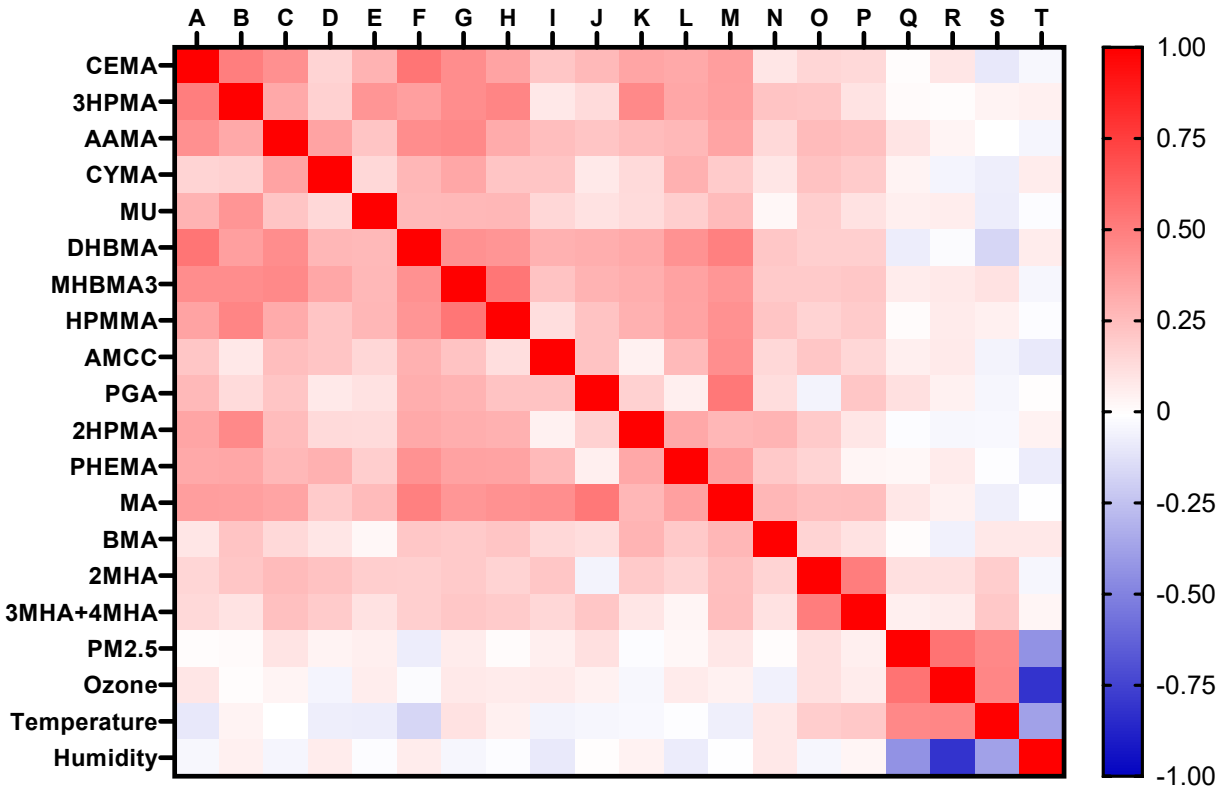
VOC metabolite	Male	Female	White	Black	Age <55	Age ≥55	BMI <30	BMI ≥30
n	135	240	311	43	202	173	199	176
CEMA	104.1 ± 87.6	122.4 ± 130.4	111.2 ± 102.5	136.1 ± 110.7	102.1 ± 78.4	131.8 ± 148.7	113.1 ± 132.6	118.9 ± 96.8
3HPMA	326.2 ± 269.2	350.2 ± 380.8	333.6 ± 313.3	330.3 ± 395.3	337.3 ± 325.8	346.5 ± 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.

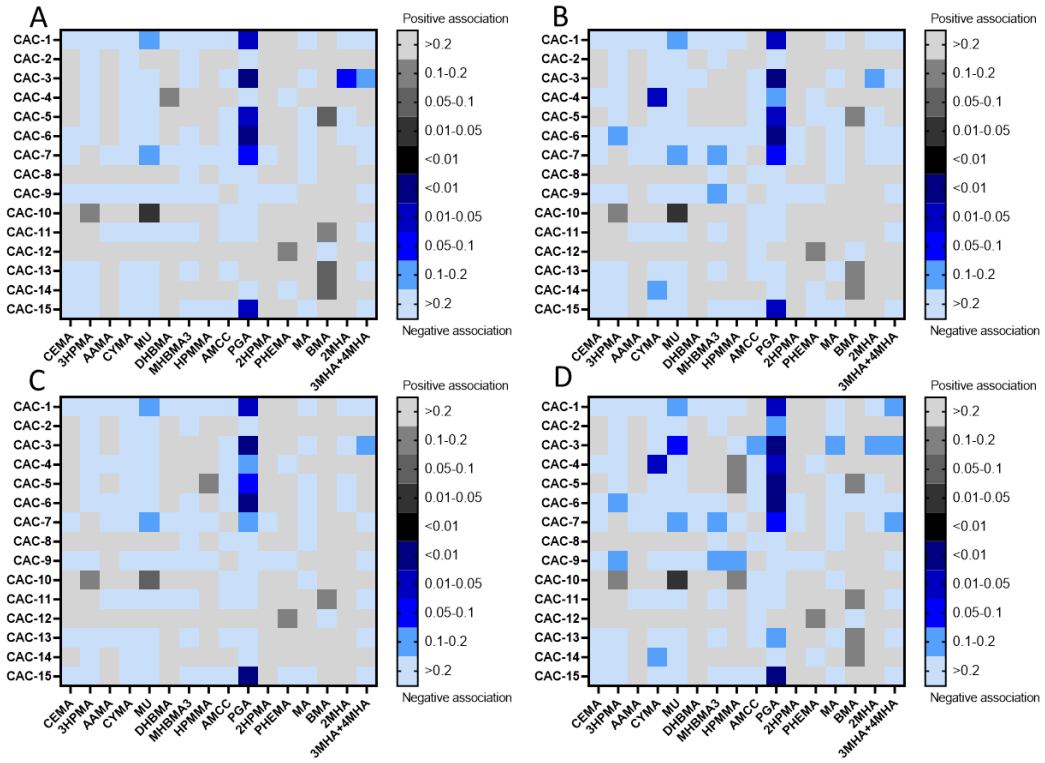
Supplemental Table 2 continued. Summary Statistics of urinary VOC metabolites for specific subgroups.

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/m^3. High $PM_{2.5}$ ranges from 8.0-17.7 μg/m^3.



Supplemental Figure 1. Pearson correlation coefficients between log-transformed urinary VOC metabolites and ambient pollutant and meteorological data (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 (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.