

1 Identification of occupations susceptible to high exposure and risk associated with multiple
2 toxicants in an observational study: National Health and Nutrition Examination Survey 1999-
3 2014

4
5 Vy Kim Nguyen Ph.D.^{1,3}, Justin Colacino Ph.D.^{1,2,4,5}, Chirag J Patel Ph.D.³, Maureen Sartor
6 Ph.D.², Olivier Jolliet Ph.D.^{1,2}

7
8 Affiliations:

9 1 Department of Environmental Health Sciences, School of Public Health, University of
10 Michigan, Ann Arbor, MI, USA

11 2 Department of Computational Medicine and Bioinformatics, Medical School, University of
12 Michigan, Ann Arbor, MI, USA

13 3 Department of Biomedical Informatics, Medical School, Harvard University, Boston, MA,
14 USA

15 4 Department of Nutritional Sciences, School of Public Health, University of Michigan, Ann
16 Arbor, MI, USA

17 5 Rogel Cancer Center, University of Michigan, Ann Arbor, MI, USA

18

19 To whom correspondence should be addressed:

20 Vy Kim Nguyen

21 Department of Biomedical Informatics, Harvard Medical School

22 10 Shattuck Street, Boston, MA, 02115, USA

23 e-mail: vy_nguyen@hms.harvard.edu

24 Running title: Occupations susceptible to high exposure and risk

25

26 **Declaration of Competing Interest**

27 At the time of the writing of the manuscript, Dr. Patel was a consultant and shareholder of

28 XY.health. Dr. Patel is also paid by J&J. All other authors declare they have nothing to disclose.

29 **Abstract**

30 **Background:** According to the World Health Organization, occupational exposures to hazardous
31 chemicals are estimated to cause over 370,000 premature annual deaths. The risks due to multiple
32 workplace chemical exposures, and those occupations most susceptible to the resulting health
33 effects, remain poorly characterized.

34 **Objectives:** The aim of this study is to identify occupations with elevated toxicant biomarker
35 concentrations and increased health risk associated with toxicant exposures in a working US
36 population from diverse categories of occupation. More specifically, we aim to 1) define
37 differences in chemical exposures based on occupation description, 2) identify occupational
38 groups with similar chemical exposure profiles, and 3) identify occupational groups with chemical
39 biomarker levels exceeding acceptable health-based biomarker levels.

40 **Methods:** For this observational study of 51,008 participants, we used data from the 1999-2014
41 National Health and Nutrition Examination Survey. We characterized differences in chemical
42 exposures by occupational group for 129 chemicals by applying a series of generalized linear
43 models with the outcome as biomarker concentrations and the main predictor as the occupational
44 groups, adjusting for age, sex, race/ethnicity, poverty income ratio, study period, and biomarker
45 of tobacco use. We identified groups of occupations with similar chemical exposure profiles via
46 hierarchical clustering. For each occupational group, we calculated percentages of participants
47 with chemical biomarker levels exceeding acceptable health-based guidelines.

48 **Results:** Blue collar workers from “Construction”, “Professional, Scientific, Technical Services”,
49 “Real Estate, Rental, Leasing”, “Manufacturing”, and “Wholesale Trade” have higher biomarker
50 levels of toxic chemicals such as several heavy metals, acrylamide, glycidamide, and several
51 volatile organic compounds compared to their white-collar counterparts. For these toxicants, 1-
52 58% of blue-collar workers from these industries have toxicant concentrations exceeding
53 acceptable levels.

54 **Discussion:** Blue collar workers have toxicant levels higher relative to their white-collar
55 counterparts, often exceeding acceptable levels associated with noncancer effects. Our findings
56 identify multiple occupations to prioritize for targeted interventions and health policies to monitor
57 and reduce high toxicant exposures.

58

59

60

61

62 **1. Introduction**

63 Data from the World Health Organization suggest that exposures to hazardous chemicals in an
64 occupational setting are responsible for over 370,000 premature deaths annually (Lim et al., 2012;
65 Stanaway et al., 2018). Such findings lend urgency to characterize occupational exposures to
66 identify workers from which industries or jobs are susceptible to adverse effects from toxicant
67 exposures. Many studies tend to focus on one chemical or one chemical family to evaluate
68 occupational exposures (Birks et al., 2016; Brenner et al., 2015). In doing so, these studies may
69 miss the complete picture of being exposed to a slew of toxicants if the focus is only directed at
70 one chemical or one chemical family. Furthermore, exposures to multiple chemicals can further
71 increase the risk of a disease. For a few studies that have investigated across multiple chemicals,
72 they have narrowed their focus on a limited set of industries and job titles (Kijko et al., 2016; Mater
73 et al., 2016). Thus, there is a need for a comprehensive, untargeted approach to study occupational
74 exposures for a wide range of chemicals across a variety of occupations.

75 Furthermore, many studies on occupational exposures use estimates of exposures based on job
76 titles or air measurements at the workplace (Birks et al., 2016; Kijko et al., 2016; Mater et al.,
77 2016). These indirect measures are limited in their ability to accurately quantify the distribution of
78 chemical exposures within the human body. In contrast, human biomonitoring provides a more
79 direct estimate of exposure while also integrating exposures which derive from multiple sources
80 and pathways. In addition, another advantage of biomonitoring data is that it provides an internal
81 dose that can be related to a toxicological response (Hays et al., 2007). In particular, biomonitoring
82 equivalents define a concentration cutoff of a chemical or its metabolites in a biological medium
83 such as blood, urine, or serum based on acceptable exposure values such as reference dose,
84 tolerable daily intakes, or minimal risk levels (Aylward et al., 2013). Several studies have used

85 biomonitoring equivalents as a screening method to evaluate risk from exposures to environmental
86 toxicants in the general population (Aylward et al., 2013; Faure et al., 2020). However, few studies
87 have used biomonitoring equivalents in an occupational context to determine prevalence of
88 workers with concentrations above acceptable levels by industry and job description (Aylward et
89 al., 2010; Hays et al., 2008). Such insight will help identify which toxicants and occupations should
90 be prioritized for further human biomonitoring, health risk evaluation, and targeted interventions.
91 Our goal is to broadly identify occupations susceptible to high exposure and risk associated with
92 combinations of multiple toxicants. To accomplish this goal, we used data from the National
93 Health and Nutrition Examination Survey (NHANES), which measures a broad range of 517
94 chemical biomarkers as part of their mission to assess the health and nutritional status of the US
95 noninstitutionalized population. Occupational information, particularly the 21 industrial and 19
96 occupation codes, are also available. Our objectives are to 1) define differences in chemical
97 exposures based on occupation description, 2) identify occupational groups with similar chemical
98 exposure profiles, and 3) identify occupational groups with chemical biomarker levels exceeding
99 acceptable health-based biomarker levels.

100 **2. Methods**

101 **2.1 Study Population**

102 Since 1999, the Centers for Disease Control (CDC) has conducted the Continuous NHANES to
103 collect cross-sectional data on demographic, socioeconomic, dietary, and health-related
104 information in the US population. For this analysis, we combined data from the chemical
105 biomarker, demographic, and occupational datasets between years 1999-2014 for an initial sample
106 of 82,091 participants. We categorized participants as different groups of unemployment status
107 using the questionnaires on the type of work done last week and main reason for not working last

108 week. We categorized the workers into their corresponding industry by using the publicly available
109 industry code on the participants' current job. We categorized participants into white- or blue-
110 collar workers by using the publicly available occupational codes on the participants' current job
111 and the US Department of Labor definition of blue-collar (U.S. Department of Labor, 2019). Blue-
112 collar workers are defined as workers who perform repetitive tasks with their hands, physical skill,
113 and energy. The industry and occupational codes can be found at
114 <https://wwwn.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=Questionnaire>. We
115 tabulated the job occupation description and the collar category in Table S1. We then excluded
116 participants under 16 years old (N = 30,987) as this is the minimum age at which NHANES
117 recorded occupation status. We excluded participants who recorded their industry as "Blank but
118 applicable" (N = 9). We also excluded participants (N = 87) from the following occupational
119 groups if the sample size was less than 50 participants: blue-collar workers from "Armed Forces"
120 and "Finance, Insurance" and white-collar workers from "Armed Forces", "Private Household",
121 and "Mining". These exclusion and inclusion criteria are further detailed in **Figure 1**. The resulting
122 sample size of our studied population was 51,008 participants. Table S2 provides the sample size
123 of each industry-collar combinations and unemployment.

124 The National Center for Health Statistics research ethics review board provided ethical approval
125 of the study. All participants provided written informed consent.

126 **2.2 Chemical Biomarkers of Occupational Exposures**

127 We defined chemical biomarker, c , as an indicator of environmental exposure that can be measured
128 in blood, serum, or urine. We replaced all measurements below the limit of detection (LOD) with
129 the LOD divided by the square root of 2, as recommended by the CDC (CDC, 2009) to produce
130 reasonably unbiased means and standard deviations (Hornung & Reed, 1990). For di-2-ethylhexyl

131 phthalate (DEHP) and arsenic separately, we calculated the sum of metabolites to compare with
132 respect to the biomonitoring equivalents. We used mono-(2-ethyl-5-oxohexyl) phthalate, mono-
133 (2-ethyl-5-hydroxyhexyl) phthalate, mono-(2-ethylhexyl) phthalate, and mono-2-ethyl-5-
134 carboxypentyl phthalate to calculate the summation of DEHP metabolites. We calculated the
135 summation of arsenic metabolites with monomethylarsonic acid and dimethylarsonic acid.
136 Therefore, we have a total of $c = 517$ chemical biomarkers.
137 Then, we further excluded chemicals that a) have a median sample size of less than 90 participants
138 across the occupational groups, b) have non-lipid adjusted measurements when lipid adjusted
139 measurements are available and c) have a detection frequency less than 10%. We delineated in
140 detail which chemical biomarkers were excluded from our analyses in Text S1. We tabulated the
141 inclusion criteria for each chemical in Table S4. The final dataset for analysis consisted of 129
142 chemical biomarkers from 12 classes (Figure 1 and Table S5). We tabulated the sample size of
143 each chemical in Table S6. We tabulated the distribution statistics for each chemical in Table S7.
144 We displayed the detection frequency by each combination of chemical biomarker and
145 occupational group in Figure S1. Laboratory methods used to measure the chemical biomarkers
146 are provided at
147 <https://wwwn.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=Laboratory>.

148 **2.3 Statistical Analysis**

149 We performed all analyses using R version 3.6.0. Our analytic code is publicly available on
150 GitHub (https://github.com/vynguyen92/nhanes_occupational_exposures). We applied the survey
151 weights to all our statistical models to 1) account for NHANES sampling designs and 2) enable
152 the generalizability of our findings to the non-institutionalized, civilian US population.

153 We used multivariate regression models to evaluate differences in the chemical biomarker levels
154 across the occupational groups, which includes the industry-collar combinations and unemployed
155 groups. We conducted a series of stepwise linear regression models with the \log_{10} transformed
156 chemical measurements as the outcome variable and the main predictor as the occupation groups
157 with the reference group as white collars from Public Administration. The selection of the
158 reference group was based on the *a priori* hypothesis that white collar workers from Public
159 Administration would be exposed to toxicants at relatively low levels for most toxicants. We
160 adjusted for age (continuous), gender (categorical), race/ethnicity (categorical), poverty income
161 ratio (continuous), NHANES cycle (continuous), and tobacco use/exposure status using serum
162 cotinine levels (continuous). Race is self-reported by the participants. The poverty income ratio is
163 defined by dividing the total family income by the poverty income line. A poverty income ratio
164 lower than 1 implies that the participant's total family income is below the poverty income line.
165 For ease of interpretation, the regression coefficients for the occupational groups were converted
166 to percent differences $[10^{\text{coefficient}} - 1] \times 100$. To identify significant comparisons while maintaining
167 a lower false positive rate, we used the False Discovery Rate (FDR) method on the p-values of the
168 regression coefficients pertaining to the occupational groups (Benjamini & Hochberg, 1995).
169 The non-random sparsity of the chemical biomarker dataset in the NHANES worker population
170 creates challenges in applying machine learning techniques to group individual workers together
171 based on similarity in chemical exposure profiles. Applying most machine learning techniques
172 requires a complete dataset (Soley-Bori, 2013). However, as no worker has data available for all
173 studied toxicants (Table S8, Figure S2), we cannot characterize the chemical profile for each
174 individual worker. Such challenges limited studies to characterizing combinations of toxicant
175 exposures within a specific chemical family, but workers are exposed to multiple toxicants across

176 a variety of chemical families at their workplace. Instead of characterizing the toxicant profile of
177 a given worker, we can characterize the profile for a group of individual workers to address this
178 sparsity issue by identifying clusters of occupations with similar chemical exposure profiles. Thus,
179 we performed hierarchical agglomerative clustering analyses on the dataset of percent differences
180 for the occupational groups. Text S2 describes the methodology used to identify clusters of
181 occupations with similar chemical exposure profiles.

182 To identify susceptible occupations with chemical biomarker levels exceeding acceptable
183 exposure levels, we compared workers' biomarker levels to biomonitoring equivalents, i.e. the
184 internal levels corresponding to these acceptable exposures. We, first, performed a literature
185 review to develop a database of biomonitoring equivalents (Table S10-S11). There are three types
186 of effects for the biomonitoring equivalents: noncancer, inhalation cancer, and ingestion cancer.
187 Noncancer effects can include mutagenicity, developmental toxicity, neurotoxicity, and
188 reproductive toxicity. Cancer effects are specific for different route of exposures. Then we used
189 this database to calculate the percentage of participants above the biomonitoring equivalent for
190 each occupation and each chemical. We used hierarchical clustering to group the occupations who
191 have similar profiles in chemical biomarker levels exceeding acceptable levels.

192 Text S3 provided full details on the methodology to quantify the contribution of occupation in
193 explaining chemical biomarker levels. Text S4 provided full details on the sensitivity analyses to
194 characterize the influence of smoking on differences in chemical biomarker levels by occupation.

195 **2.4 Role of the funding source**

196 The funders of the study did not have a role in study design, data collection, data analysis, data
197 interpretation, and writing of the manuscript. All authors have full access to the data in the study

198 and accept responsibility for the decision to submit for publication. The corresponding author had
199 full access to all the data and the final responsibility to submit for publication.

200 **3. Results**

201 All figures are available on our interactive app at
202 https://chiragip.shinyapps.io/nhanes_occupational_exposures/.

203 **3.1 Study Population**

204 **Table 1 and 2** presents population characteristics for the 51,008 NHANES participants from 1999-
205 2014. **Figure 2** shows the percentage of categories for age group, sex, race, poverty income ratio,
206 smoking status, and study period for each occupational group. Participants working in blue-collar
207 jobs tend to be on average younger compared to those working in white-collar jobs (**Figure 2A**).
208 Blue-collar jobs are primarily occupied by males, while females tend to work in private household,
209 health care, and education (**Figure 2B**). White-collar jobs are predominantly comprised of Non-
210 Hispanic White participants, whereas blue-collar workers tend to be more diverse (**Figure 2C**).
211 There is a socioeconomic gradient with white-collar workers having higher poverty income ratio
212 (i.e. lower socioeconomic status) compared to blue-collar workers (**Figure 2D**). There is a
213 substantial proportion of active smokers in blue collar jobs as well as those “Looking for work”,
214 “Disabled”, or “On layoff”, whereas the proportions of active smokers are much lower in white-
215 collar jobs (**Figure 2E**). **There is** a relative uniform distribution of participants by study period for
216 each occupation (**Figure 2F**). These figures are available on our interactive app at
217 https://chiragip.shinyapps.io/nhanes_occupational_exposures/.

218 **3.2 Differences in Chemical Biomarker Levels by Occupational Groups**

219 **Figure 3** displays the differences in chemical biomarker levels across the occupational groups
220 using regression and distribution statistics for the following chemicals: lead, m-/p-xylene,

221 cotinine, 2,4-D, glycidamide, and sum of DEHP metabolites. We selected these chemicals for the
222 at least one of the following reasons: 1) availability of having a biomonitoring equivalent and 2)
223 existence of statistically significant differences in biomarker levels across the occupational groups.
224 Biomonitoring equivalents are available for the selected chemicals except for cotinine. In Figure
225 3A, blood lead is, on average, significantly higher in most blue-collar workers and unemployed
226 groups such as “Looking for work”, “On layoff”, “Disabled”, and “Retired” compared to the other
227 white-collar workers. We observed a similar pattern in blood cadmium (Figure S4). In contrast,
228 metabolites of mercury display the opposite exposure patterns to those of lead and cadmium. Total
229 blood mercury levels of most blue-collar workers are substantially and significantly lower
230 compared to those of white-collar workers (Figure S5). Similarly, m-/p-xylene (Figure 3B) and
231 toluene (Figure S6) are higher in blue-collars and unemployed participants in “On layoff”,
232 “Disabled”, and “Unable to work for health reasons” compared to white-collar workers. Similar
233 results are observed for several Polyaromatic Hydrocarbons (PAHs) such as 1-pyrene, 2-fluorene,
234 3-fluorene, and 1-naphthol (Figure S7), but the signals are not as strong as those for toluene and
235 m-/p-xylene. It is noteworthy that within the same industry such as “Professional, Scientific,
236 Technical Services”, levels of m-/p-xylene are substantial different between white versus blue
237 collars. Figure 3C shows a smoking gradient with blue-collar workers and unemployed participants
238 having substantially higher levels of cotinine compared to white collars and the NHANES
239 populations. NNAL, which is primarily found in tobacco products, shows a similar trend (Figure
240 S8). The signals for the smoking related compounds are among the strongest and most substantial.
241 Figure 4D shows how concentrations of a herbicide, 2,4-D, are significantly and substantially
242 higher in blue-collar and white-collar workers from “Agriculture, Forestry, Fishing” along with
243 blue-collar workers from “Information Services”. Though, none of the participants have 2,4-D

244 levels exceeding the guideline biomonitoring equivalent levels. DEET acid, a metabolite of DEET
245 and a common ingredient used in insect repellent, shows a similar pattern (Figure S9). In **Figure**
246 **3E**, glycidamide levels, on average, are significantly higher in blue collars from “Wholesale
247 Trade”, “Other Services”, “Professional, Scientific, Technical Services”, “Retail Trade”,
248 “Construction”, and “Arts, Entertainment, Recreation”. In addition, glycidamide levels are also
249 significantly higher in white collars from “Accommodation, Food Services”, “Agriculture,
250 Forestry, Fishing”, and “Retail Trade”. Acrylamide show similar pattern (Figure S10). These two
251 chemicals are found in food prepared at high temperature via frying, baking, or roasting and are
252 used in textile, paper processing, and cosmetics. In **Figure 3F**, levels of the sum of urinary DEHP
253 metabolites are significantly higher in “Professional, Scientific, Technical Services” and
254 “Accommodation, Food Services” compared to the reference group. Phthalates in general are used
255 as plasticizers. On the contrary, mono-ethyl phthalate, an indicator of personal care product usage,
256 are, on average higher, in the reference group (Figure S11).

257 **3.3 Chemical Exposure Profiles**

258 **Figure 4** shows differences in chemical exposure profiles by occupational groups for the 129
259 studied chemicals. The chemical exposure profiles of blue-collar workers are more similar to those
260 of other blue-collar workers than to their white-collar counterparts. For example, blue-collar
261 workers from “Construction”, “Other Services”, “Professional, Scientific, Technical Services”,
262 “Real Estate, Rental, Leasing”, “Manufacturing”, and “Wholesale Trade” have some of the highest
263 biomarker levels of heavy metals, such as cadmium and lead, PAHs, and volatile organic chemicals
264 (VOCs), including m-/p-xylene and toluene, but have lower levels of arsenic and mercury
265 metabolites and BP-3, which is a UV blocking chemical used in sunscreen. There are some blue
266 collars who have higher levels of dietary components such as orange or red plant pigments found

267 in fruits and vegetables such as trans- β -carotene, cis- β -carotene, and α -carotene and a form of
268 vitamin E such as d-Tocopherol. Interestingly, participants who are “Looking for work”, “On
269 layoff”, “Disabled”, “Unable to work for health reasons”, and “Occupation Missing” have similar
270 chemical exposure profiles to the aforementioned blue-collar workers. Participants who report
271 being “Disabled” or “Unable to work for health reasons” have significantly lower levels of several
272 dietary components. Within the food services cluster, which includes blue and white collars from
273 “Transportation, Warehousing” and “Accommodation, Food Services”, the phthalates signal is
274 particularly stronger in the blue-collar workers from “Accommodation, Food Services”. The far
275 left of the heatmap consists of mostly white-collar workers. These occupational groups have the
276 most similar chemical exposure profiles to that of white-collar workers from “Public
277 Administration” as the percent differences across most studied chemicals are near 0, i.e. the blue
278 and red boxes are faded. We excluded smoking-related compounds such as NNAL and Cotinine
279 from the figure, or else we would not be able to observe any signals from the other chemical
280 biomarkers (Figure S12).

281 **3.4 Occupational Groups with Chemical Biomarker Levels Exceeding Acceptable Guidelines**

282 **Figure 5** shows the percentage of a given occupational group with chemical biomarker levels
283 exceeding the biomonitoring equivalents for noncancer effects. Our hierarchical clustering
284 analysis on the occupational groups show that the left half of Figure 5 includes predominantly blue
285 collars and unemployed groups, while the right half includes predominantly white collars. This
286 suggests that blue collars and unemployed groups have similar chemical exposure profiles to each
287 other, and such profiles are different from the chemical profiles of the white collars. Several blue
288 collar jobs along with unemployed groups such as those who are “On layoff”, “Unable to work for
289 health reasons”, and “Disabled” have some of the highest percentages of participants with

290 biomarker levels exceeding acceptable health levels for VOCs such as m-/p-xylene, benzene,
291 pesticides such as 3-phenoxybenzoic acid, heavy metals such as cadmium, lead, and arsenic,
292 metabolites of DEHP, and acrylamide and its metabolite glycidamide. These findings suggest that
293 these toxicants for the aforementioned occupations may be further monitored to understand why
294 biomarker levels are exceeding acceptable guidelines. Figure S13-S14 shows the percentage of
295 participants with excessive toxicant levels for the cancer effects.

296 **4. Discussion**

297 In this study, we systematically characterize differences in chemical biomarker levels across a
298 diverse suite of chemical contaminants and occupations. This is the first application of hierarchical
299 clustering on differences by chemical exposures to identify groups of workers with similar
300 chemical exposure profiles. This is also the first study to determine the percentage of a given
301 occupation who are exceeding acceptable levels for a broad set of toxicants. Furthermore, this is
302 the first application of hierarchical clustering to systematically identify which chemicals for which
303 occupations have biomarker levels exceeding acceptable guidelines. Our findings are informative
304 for identifying which workers are susceptible to higher exposures from which toxicants.

305 Contact with products and equipment may explained higher biomarker levels of heavy metals such
306 as lead and cadmium, PAHs, and VOCs such as toluene and benzene found in blue-collar workers.
307 Higher lead levels found in blue collars may be due to the presence of lead in old and commercial
308 paint, car parts, batteries, glass, and consumer products made of plastics (CDC, 2018). In addition,
309 this same group of workers may be exposed to cadmium via industrial uses of cadmium in making
310 batteries, plating, pigments, and plastics (Miao & Ji, 2019). Sources of occupational PAH
311 exposures to this group may be due to engaging in tasks that involve combustion emission
312 (Raymond, 1998). Similarly, higher VOCs levels may also be due to working with products

313 containing VOCs (Minnesota Department of Health, 2020). Overall, higher biomarker levels of
314 heavy metals, PAHs, and VOCs in predominantly blue-collar workers may be due to contact with
315 products containing these chemicals. This suggests that the forementioned blue collars should be
316 further examined to understand sources of exposures for such toxicants and reasons for exceeding
317 acceptable levels.

318 On the other hand, behaviors associated with higher socioeconomic status may explain why most
319 white-collar workers have higher levels of metabolites of arsenic and mercury along with a
320 biomarker of sunscreen use, BP-3. While arsenic (CDC, 2019a) and mercury (CDC, 2019b) are
321 used in many industries, it is less likely that higher biomarker levels of arsenic metabolites in white
322 collars are due to occupational exposures. Although, health care workers may be exposed to
323 mercury via medical or dental equipment (Trzcinka-Ochocka et al., 2007). Higher mercury and
324 arsenic biomarker levels among these white-collar workers may indicate higher fish consumption
325 (Shimshack et al., 2007), which associated with higher socioeconomic status instead of an
326 indicator of occupational exposures (Tyrrell et al., 2013). It is also doubtful that white-collar
327 workers are manufacturing products containing exposed to BP-3 (Benzophenone, 2013; Program,
328 2006). Instead, as BP-3 is used to prevent UV light from damaging scents and colors in personal
329 care products (CDC, 2017), it more likely that higher levels of this chemical may suggest that
330 cosmetics usage has a major role in strategic self-presentation. Overall, the chemical exposure
331 profiles of white-collar workers likely indicate behaviors associated with socioeconomic status.

332 Many studies using NHANES have been limited to studying chemical co-exposures in one
333 chemical family due to the non-random sparsity of the chemical biomarker data (Sadetzki et al.,
334 2016; Shim et al., 2017). To address this sparsity challenge, we conducted clustering analysis on
335 exposure differences among the occupational groups, i.e. clustering analysis on statistics of the

336 biomarker data instead of on the raw data. Our framework enabled the identification of co-
337 exposure across a wide range of chemicals not only limited to one chemical family. This
338 framework can be applied in other settings to help cluster observations based on similar profiles
339 especially in a non-randomly sparse dataset. This can be done without having to form a complete
340 dataset or impute the missing values.

341 Chemical exposures have been implicated as etiological agents in adverse noncancer effects on the
342 nervous, reproductive, immune, and cardiovascular systems. Such toxic chemicals include heavy
343 metals such as lead and cadmium, VOCs, PAHs, phthalates, and smoking-related compounds.
344 Several epidemiological and toxicological studies have implicated the neurotoxicity of heavy
345 metals, especially lead (Garza et al., 2006) and cadmium (Méndez-Armenta & Ríos, 2007). Many
346 studies have shown the reproductive toxicity of phthalates (Martino-Andrade & Chahoud, 2010)
347 and PAHs (Yin et al., 2017). PAHs (Schober et al., 2007) and VOCs such as toluene and benzene
348 (Jiménez-Garza et al., 2017) are also known to elicit an inflammatory response individually and
349 in combination (Kuang et al., 2020). Heavy metals have been recently linked to cardiovascular
350 disease (Duan et al., 2020). Tobacco exposures have been causal factors in several noncancer
351 effects such as respiratory problems (Anderson & Ferris Jr, 1962), heart disease (Lakier, 1992),
352 infections (Arcavi & Benowitz, 2004), and fertility problems (Stillman et al., 1986). Overall, our
353 findings of toxicant levels in blue collar workers exceeding acceptable levels coupled with existing
354 literature on the adverse effects of such toxicants continues to further implicate blue collars from
355 the aforementioned industries to be a high-risk population.

356 Our findings lend urgency to understand how one's occupation can be a route for smoking
357 initiation in young people and consequentially exposure to other toxicants. The connection
358 between being a blue collar, active smoker, and younger may suggest that being an active smoker

359 is part of the culture of a blue-collar worker (Ham et al., 2011). Furthermore, blue-collar workers
360 are additionally exposed to VOCs, PAHs, and heavy metals, since these chemicals have been
361 detected in tobacco products. Interestingly, blue-collar workers have higher levels of nicotine
362 metabolites and dietary components such as beta- and alpha-carotenes. This finding is especially
363 alarming, since cancer-preventative trials have shown that vitamin A analogues, alone or in
364 combination with vitamin E, are risk factors for lung cancer and mortality in active smokers
365 (Omenn et al., 1996; Paolini et al., 1999). Our findings can inspire future studies to develop
366 interventions to understand cultural and behavioral factors leading to smoking initialization and
367 implement evidence-based regulations on tobacco control to prevent the younger population from
368 initiating smoking (Kunst, 2021).

369 Environmental injustice is defined as the disproportionate exposures of toxicants and their
370 consequential effects on health to disadvantaged groups such as individuals from racial minorities
371 and/or low socioeconomic status (Maantay, 2002). Blue-collar workers are disproportionately
372 exposed to some of the most toxic chemicals with several found at concentration exceeding
373 acceptable guidelines for cancer and noncancer effects (Kolonel, 1976; Swartz, 2001).
374 Furthermore, many blue-collar workers may come from the lowest socioeconomic status
375 (Kivimäki et al., 2006). Thus, our findings call for an increase in exposure surveillance and
376 industrial controls, effect regulations on chemical exposures, inform remediation strategies, and
377 help implement interventions programs to improve the health of workers susceptible to toxicant
378 exposures. Moreover, health providers such as occupational physicians and industrial health
379 personnel can use our findings to inform their patients on preventatives measures to avoid
380 occupations and toxicant exposures that may increase their personal and familial disease risk.

381 The present study has several other limitations. First, as NHANES is a cross-sectional survey, we
382 cannot make claims on causal factors of chemical exposures. Second, a limitation of using
383 chemical biomarker data is that a delay between the time of exposure and time of data collection
384 may prevent the detection of higher occupational exposures. This limitation is especially salient
385 for VOCs, which have short half-life ranging from 2 to 128 hours (OECD, 2004, 2014), which
386 implies that substantially higher biomarker levels could be observed at workplace. Third, while
387 we identified differences in chemical biomarker levels by occupation, we cannot claim that such
388 differences are due to occupational exposures, since we do not know the source of all exposures
389 for each study participant. Nevertheless, our findings can inspire future studies to prioritize
390 chemicals and susceptible occupations in specific industries to measure at the workplace. Fourth,
391 while NHANES has biomonitoring data on 517 chemicals, we only obtained biomonitoring
392 equivalents for 106 chemicals (20.5%). Fifth, there is potential uncertainty in how the
393 biomonitoring equivalents are converted from exposures values to biomarker concentrations. Such
394 uncertainty can either decrease or increase the value of the biomonitoring equivalent and in turn
395 change the interpretation of which chemicals for which occupations should be further monitored.

396 **5. Conclusions**

397 Evaluating differences in chemical exposures by occupation is essential to identify occupations
398 susceptible to high exposures to toxicants as well as understand how occupational exposures play
399 a role in adverse health outcomes. We applied an unbiased approach to screen across 129 chemical
400 biomarkers to characterize the chemical exposure profiles across white- and blue-collar workers
401 from 20 different industries and 7 unemployed groups. We developed a framework using
402 hierarchical clustering on differences of chemical biomarker levels to identify clusters of
403 occupations with similar chemical exposures. Our framework enabled 1) comprehensive

404 characterization of chemical exposures across a wide variety of occupations and toxicants and 2)
405 identification of occupations susceptible to high toxicant exposure and exceeding acceptable
406 health-based levels. These findings can guide efforts to design targeted interventions to reduce and
407 prevent exposures in susceptible occupations and help mitigate negative effects from toxicant
408 exposures.
409

410 **CRedit authorship contribution statement**

411 VKN: conceptualisation, data curation, formal analysis, software, visualization, writing – original

412 draft, writing – review & editing, verification of underlying data

413 JC: conceptualisation, funding acquisition, resources, writing – original draft, writing – review &

414 editing

415 CJP: funding acquisition, resources, writing – original draft, writing – review & editing

416 MS: writing – original draft, writing – review & editing

417 OJ: conceptualisation, resources, writing – original draft, writing – review & editing

418 Funding: Ravitz Family Foundation, University of Michigan Forbes Institute for Cancer

419 Discovery, Harvard Data Science Initiative, University of Michigan Center for Occupational

420 Health and Safety Engineering, and National Institutes of Health (R01 ES028802, P30 ES017885,

421 P30 CA046592, UG3 CA267907)

422 **Data Sharing**

423 Data for the study was collected by the Centers of Disease Control and Prevention and is publicly

424 available at <https://www.cdc.gov/nchs/nhanes/Default.aspx>. Data includes individual participant

425 data with corresponding demographics, occupational codes, and measurements for chemical

426 biomarkers. A data dictionary defining each field in the set is publicly available at

427 <https://www.cdc.gov/nchs/nhanes/Search/variablelist.aspx?Component=Demographics>,

428 <https://www.cdc.gov/nchs/nhanes/search/datapage.aspx?Component=Questionnaire>, and

429 <https://www.cdc.gov/nchs/nhanes/Search/variablelist.aspx?Component=Laboratory>. Our

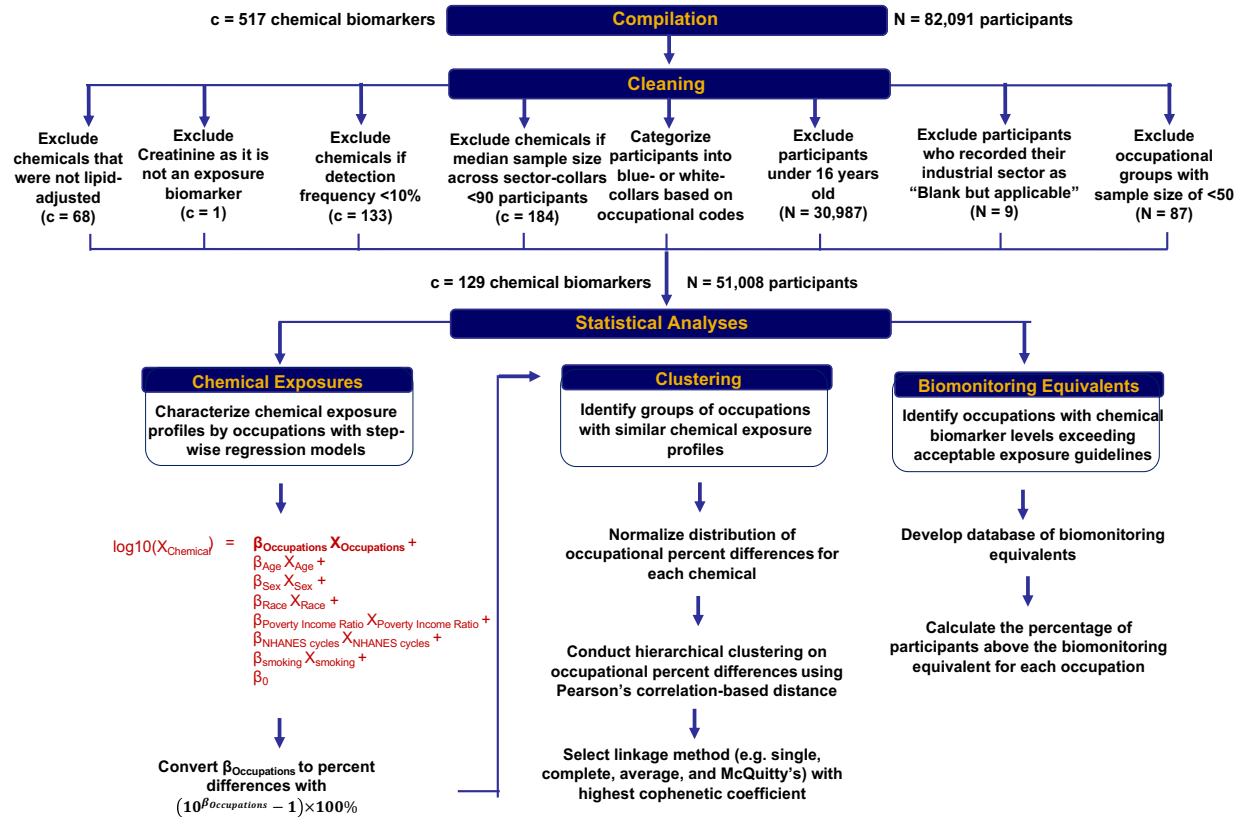
430 curated data along with related documents such as excel documents and R code for the statistical

431 analysis will be available at https://github.com/vynguyen92/nhanes_occupational_exposures with

432 publication. Data will be available to the requester via email after the approval of the request.

433

434



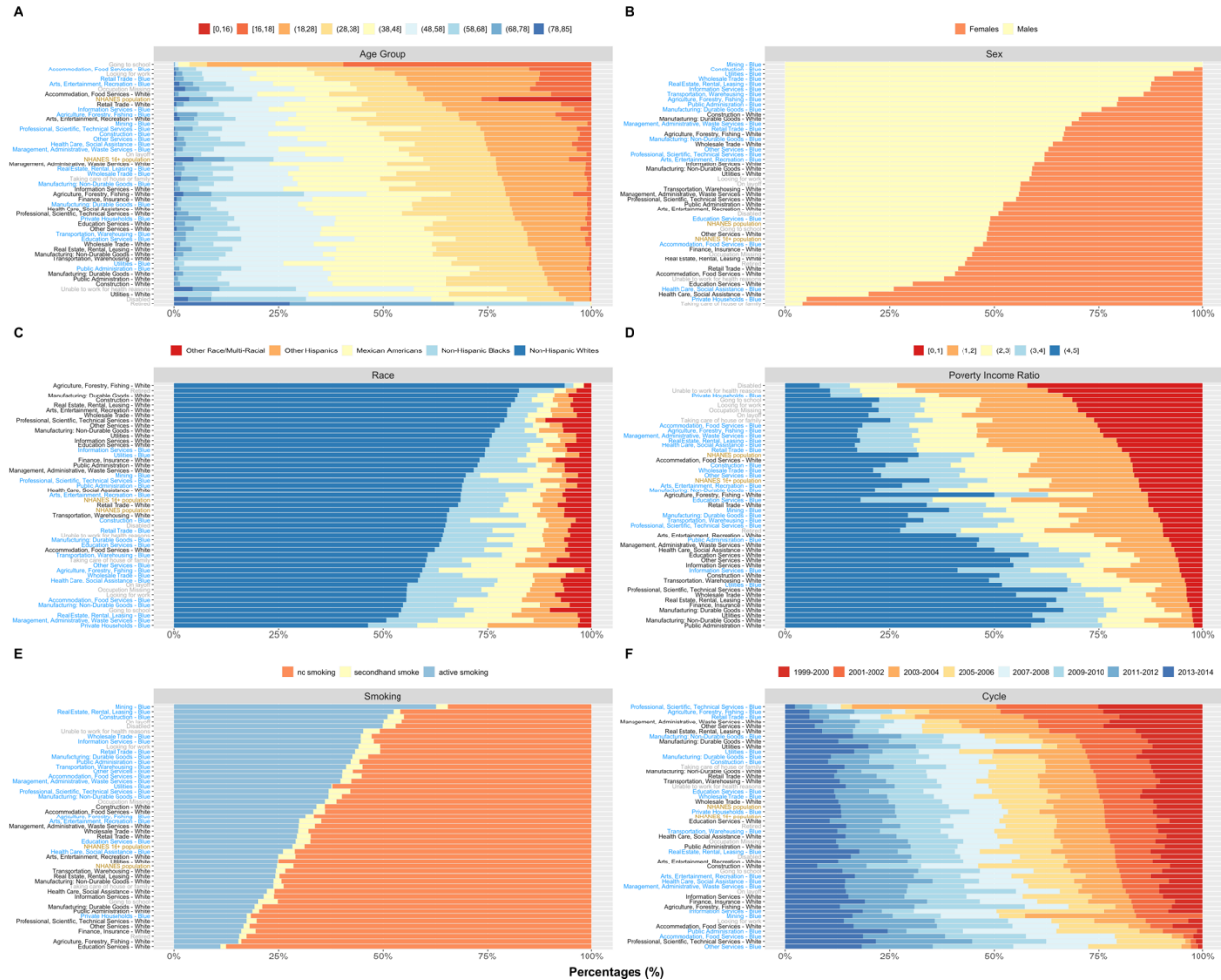
435

436 Figure 1. Schematic description on curation of chemical biomarker and inclusion criteria of
 437 participants and of the analytical methods used to characterize occupational variations in chemical
 438 exposures. Reference group for the analysis on the industry-collar combinations is white collars
 439 from public administration.

440

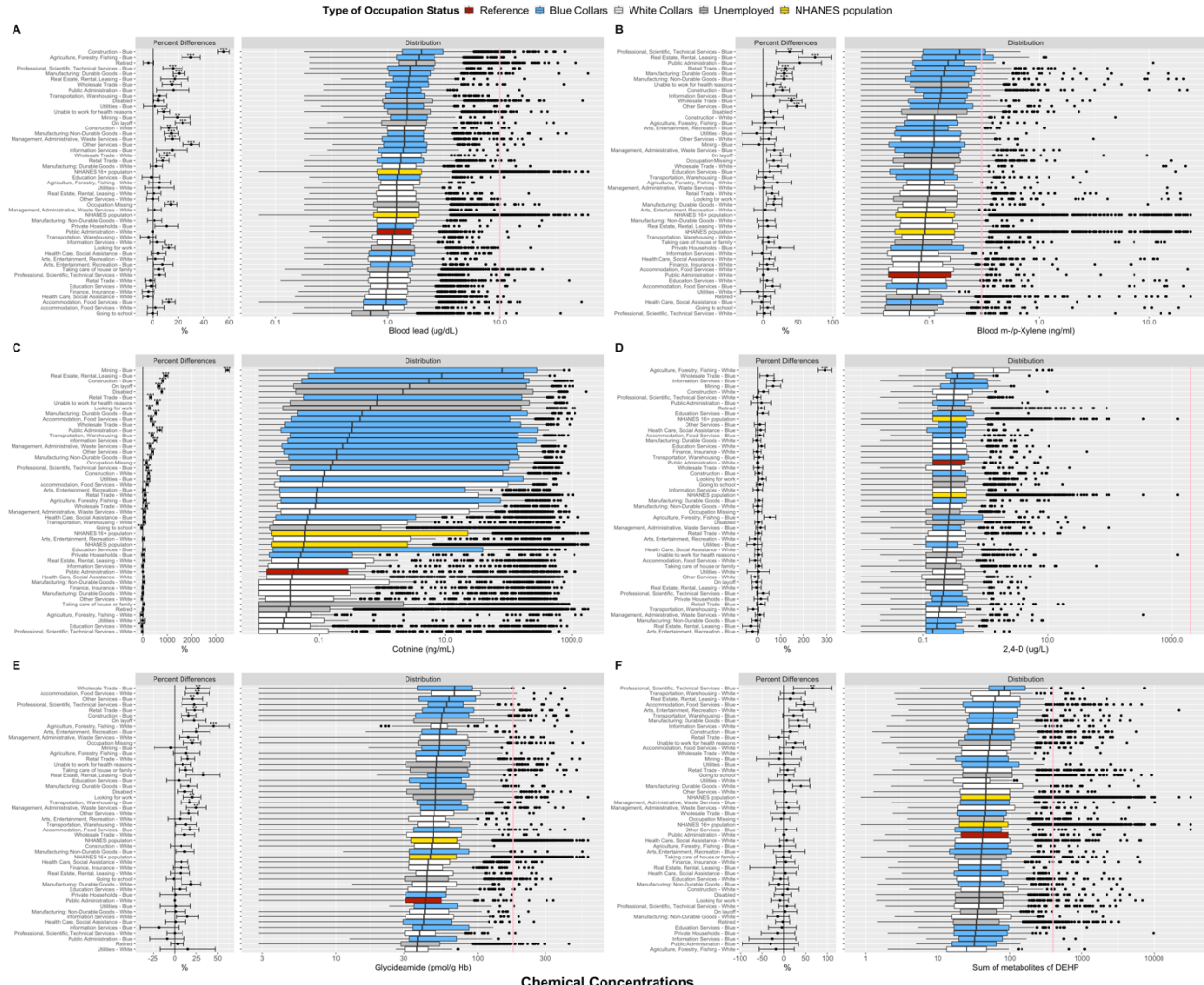
441

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



442
 443 Figure 2. Panel of bar plots showing the percentage of participants by A) age group, B) sex, C)
 444 race/ethnicity, D) poverty income ratio (PIR), E) smoking status, and F) NHANES cycle for each
 445 industry-collared combination and unemployment status. The occupational groups are ordered in
 446 ascending order based on percentage of A) participants who are 28 years and younger, B) males,
 447 C) Non-Hispanic Whites, D) PIR = [0,1] (i.e. participants who are below the poverty income line),
 448 E) participants who do not smoke, and F) participants in 1999-2002. The “NHANES Population”
 449 consists of all participants in 1999-2014. The “NHANES 16+ Population” consists of participants
 450 in 1999-2014 and are 16 years old or older. Smoking status is defined using serum cotinine levels:
 451 no smoking ≤ 1 ng/mL, secondhand smoke 1-3 ng/mL, and active smoking > 3 ng/mL. These
 452 individual figures are available on our interactive app at
 453 https://chiragip.shinyapps.io/nhanes_occupational_exposures/.
 454

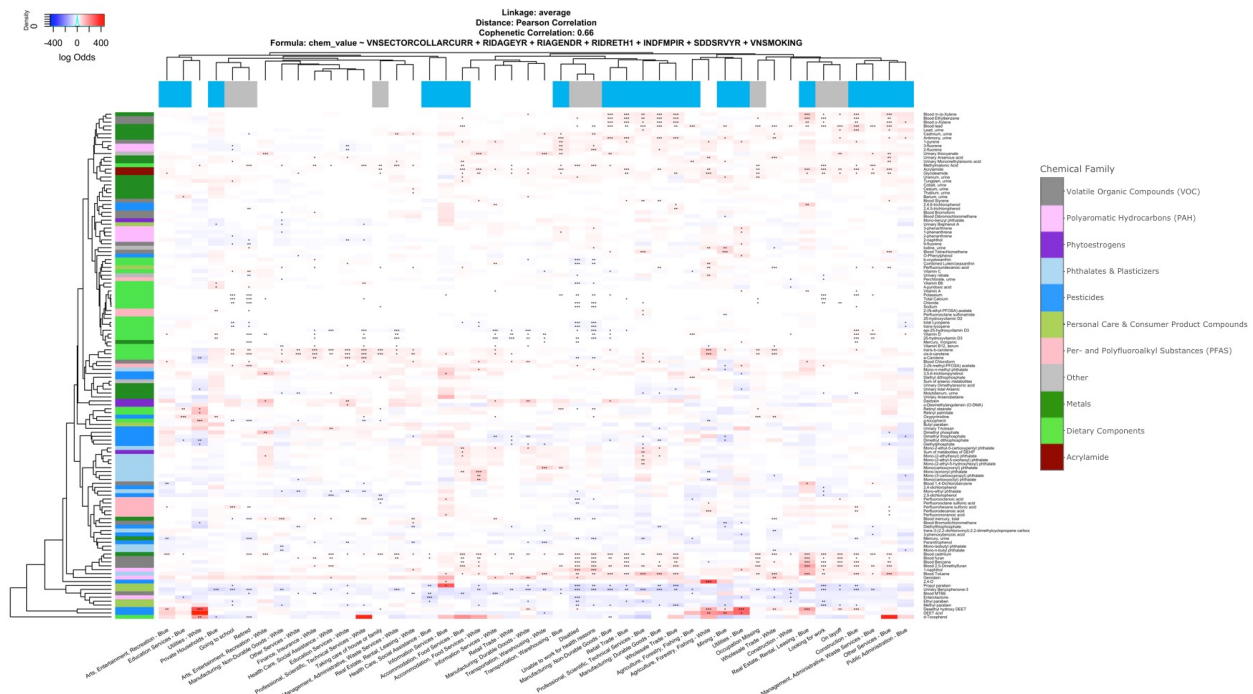
It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).



455
 456 Figure 3. Panel of boxplots of chemical distribution for A) blood Lead, B) m/p-Xylene, C)
 457 Cotinine, D) 2,4-D, E) Glycidamide, and F) sum of DEHP metabolites. Purple triangle represents
 458 the geometric mean of chemical levels for a given occupational group. The pink line represents
 459 the biomonitoring equivalent of the chemical for noncancer effects. The “NHANES Population”
 460 consists of participants in 1999-2014. The “NHANES 16+ Population” consists of participants in
 461 1999-2014 and are 16 years old or older. Percent differences are derived from fully adjusted
 462 models, which were adjusted for age, sex, race/ethnicity, poverty income ratio, study period, and
 463 serum cotinine (biomarker of smoking). Reference group for the occupational groups is comprised
 464 of white collars from Public Administration. Number of asterisks indicate statistical significance
 465 of the percent differences: * (p-value \in (0.01, 0.05]), ** (p-value \in (0.001, 0.01]), and *** (p-
 466 value \leq 0.001). The p-values corrected for multiple comparison with the Benjamini and Hochberg
 467 FDR procedure of 5%. These individual figures are available on our interactive app at
 468 https://chiragip.shinyapps.io/nhanes_occupational_exposures/.

469
 470
 471
 472

473



474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

Figure 4. Heatmap of percent differences in chemical biomarker concentrations by occupational group, relative to white collars from Public Administration. Chemical biomarkers in white color indicates that the concentrations are the same between the given industry-collar combination and the reference group. The color bar for the columns represents the collar categorization and unemployment. The color bar for the rows represents the chemical classes. Blue presents the blue-collar workers. White represents the white-collar workers. Gray presents the unemployed participants. The dendrogram of the occupational groups is defined based on using the average linkage function with Pearson's correlation-based distance. Results are adjusted for age, sex, race/ethnicity, poverty income ratio, study period, and serum cotinine (biomarker of smoking). Number of asterisks indicate statistical significance of the percent differences: * (p-value \in (0.01, 0.05]), ** (p-value \in (0.001, 0.01]), and *** (p-value \leq 0.001). This figure is available on our interactive app at https://chiragjp.shinyapps.io/nhanes_occupational_exposures/.

499

500

501 Table 1. Population statistics of the categorical variables for 51,008 NHANES participants who
502 are eligible to have an occupation title. NHANES sampling design is accounted in calculating
503 percentages (%), while counts (N) pertains to the number of NHANES participants.

	N (%)
Sex	
Males	24723 (48.2)
Females	26285 (51.8)
Race	
Mexican	10049 (8.28)
Other Hispanics	3712 (5.42)
Non-Hispanic Whites	22424 (68.59)
Non-Hispanic Blacks	11158 (11.50)
Other Race/Multi-Racial	3665 (6.21)
Cycle	
1999-2000 (Cycle 1)	6036 (10.93)
2001-2002 (Cycle 2)	6627 (12.53)
2003-2004 (Cycle 3)	6191 (12.10)
2005-2006 (Cycle 4)	6132 (12.38)
2007-2008 (Cycle 5)	6530 (12.65)
2009-2010 (Cycle 6)	6875 (12.86)
2011-2012 (Cycle 7)	6164 (13.14)
2013-2014 (Cycle 8)	6453 (13.41)

504

505

506

507 Table 2. Distribution statistics of the categorical variables for 51,008 NHANES participants who
508 are eligible to have an occupation title. NHANES sampling design is accounted in the calculations
509 of the distribution statistics, while counts (N) pertains to the number of NHANES participants.

510

	N (%)	Minimum	5 th	10 th	Median	Mean (SE)	90 th	99 th	Maximum
Age (years)	51008 (100)	16	17	21	44	44.7 (0.20)	71	83	85
Poverty Income Ratio (-)	46441 (91.0)	0	0.49	0.77	2.86	2.93 (0.031)	5	5	5
Serum cotinine (ng/mL)	45376 (88.9)	0.011	0.011	0.011	0.061	58.55 (1.56)	250	514	1820

511

512

- 513 Anderson, D. O., & Ferris Jr, B. G. (1962). Role of tobacco smoking in the causation of chronic
514 respiratory disease. *New England Journal of Medicine*, 267(16), 787–794.
- 515 Arcavi, L., & Benowitz, N. L. (2004). Cigarette smoking and infection. *Archives of Internal*
516 *Medicine*, 164(20), 2206–2216.
- 517 Aylward, L. L., Kirman, C. R., Schoeny, R., Portier, C. J., & Hays, S. M. (2013). Evaluation of
518 biomonitoring data from the CDC National Exposure Report in a risk assessment context:
519 perspectives across chemicals. *Environmental Health Perspectives*, 121(3), 287–294.
- 520 Aylward, L. L., Morgan, M. K., Arbuckle, T. E., Barr, D. B., Burns, C. J., Alexander, B. H., &
521 Hays, S. M. (2010). Biomonitoring data for 2, 4-dichlorophenoxyacetic acid in the United
522 States and Canada: interpretation in a public health risk assessment context using
523 Biomonitoring Equivalents. *Environmental Health Perspectives*, 118(2), 177–181.
- 524 Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and
525 Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B*
526 *(Methodological)*, 57(1), 289–300. <http://www.jstor.org/stable/2346101>
- 527 Benzophenone, H. (2013). *Hazardous substance database*.
- 528 Birks, L., Casas, M., Garcia, A. M., Alexander, J., Barros, H., Bergström, A., Bonde, J. P., Burdorf,
529 A., Costet, N., & Danileviciute, A. (2016). Occupational exposure to endocrine-disrupting
530 chemicals and birth weight and length of gestation: A European meta-analysis.
531 *Environmental Health Perspectives*, 124(11), 1785–1793.
- 532 Brenner, S. A., Neu-Baker, N. M., Caglayan, C., & Zurbenko, I. G. (2015). Occupational exposure
533 to airborne nanomaterials: an assessment of worker exposure to aerosolized metal oxide
534 nanoparticles in semiconductor wastewater treatment. *Journal of Occupational and*
535 *Environmental Hygiene*, 12(7), 469–481.

- 536 CDC. (2009). *Fourth National Report on Human Exposure to Environmental Chemicals*.
537 <https://www.cdc.gov/exposurereport/pdf/fourthreport.pdf>
- 538 CDC. (2017). *Benzophenone-3 (BP-3) Factsheet*.
- 539 CDC. (2018). *LEAD: Jobs That May Have Lead Exposure*.
540 <https://www.cdc.gov/niosh/topics/lead/jobs.html>
- 541 CDC. (2019a). *Arsenic*. <https://www.cdc.gov/niosh/topics/arsenic/default.html>
- 542 CDC. (2019b). *Mercury*.
- 543 Duan, W., Xu, C., Liu, Q., Xu, J., Weng, Z., Zhang, X., Basnet, T. B., Dahal, M., & Gu, A. (2020).
544 Levels of a mixture of heavy metals in blood and urine and all-cause, cardiovascular disease
545 and cancer mortality: A population-based cohort study. *Environmental Pollution*, 114630.
- 546 Faure, S., Noisel, N., Werry, K., Karthikeyan, S., Aylward, L. L., & St-Amand, A. (2020).
547 Evaluation of human biomonitoring data in a health risk based context: An updated analysis
548 of population level data from the Canadian Health Measures Survey. *International Journal*
549 *of Hygiene and Environmental Health*, 223(1), 267–280.
- 550 Garza, A., Vega, R., & Soto, E. (2006). Cellular mechanisms of lead neurotoxicity. *Medical*
551 *Science Monitor*, 12(3), RA57–RA65.
- 552 Gordon, T., Sheppard, D., McDonald, D. M., Distefano, S., & Scypinski, L. (1985). Airway
553 hyperresponsiveness and inflammation induced by toluene diisocyanate in guinea pigs.
554 *American Review of Respiratory Disease*, 132(5), 1106–1112.
- 555 Ham, D. C., Przybeck, T., Strickland, J. R., Luke, D. A., Bierut, L. J., & Evanoff, B. A. (2011).
556 Occupation and workplace policies predict smoking behaviors: analysis of national data from
557 the current population survey. *Journal of Occupational and Environmental*
558 *Medicine/American College of Occupational and Environmental Medicine*, 53(11), 1337.

- 559 Hays, S. M., Becker, R. A., Leung, H.-W., Aylward, L. L., & Pyatt, D. W. (2007). Biomonitoring
560 equivalents: a screening approach for interpreting biomonitoring results from a public health
561 risk perspective. *Regulatory Toxicology and Pharmacology*, *47*(1), 96–109.
- 562 Hays, S. M., Nordberg, M., Yager, J. W., & Aylward, L. L. (2008). Biomonitoring equivalents
563 (BE) dossier for cadmium (Cd)(CAS No. 7440-43-9). *Regulatory Toxicology and*
564 *Pharmacology*, *51*(3), S49–S56.
- 565 Hornung, R. W., & Reed, L. D. (1990). Estimation of Average Concentration in the Presence of
566 Nondetectable Values. *Applied Occupational and Environmental Hygiene*, *5*(1), 46–51.
567 <https://doi.org/10.1080/1047322X.1990.10389587>
- 568 Jiménez-Garza, O., Guo, L., Byun, H.-M., Carrieri, M., Bartolucci, G. B., Zhong, J., & Baccarelli,
569 A. A. (2017). Promoter methylation status in genes related with inflammation, nitrosative
570 stress and xenobiotic metabolism in low-level benzene exposure: Searching for biomarkers
571 of oncogenesis. *Food and Chemical Toxicology*, *109*, 669–676.
- 572 Kijko, G., Jolliet, O., & Margni, M. (2016). Occupational Health Impacts Due to Exposure to
573 Organic Chemicals over an Entire Product Life Cycle. *Environmental Science & Technology*,
574 *50*(23), 13105–13114. <https://doi.org/10.1021/acs.est.6b04434>
- 575 Kivimäki, M., Head, J., Ferrie, J. E., Brunner, E., Marmot, M. G., Vahtera, J., & Shipley, M. J.
576 (2006). Why is evidence on job strain and coronary heart disease mixed? An illustration of
577 measurement challenges in the Whitehall II study. *Psychosomatic Medicine*, *68*(3), 398–401.
- 578 Kolonel, L. N. (1976). Association of cadmium with renal cancer. *Cancer*, *37*(4), 1782–1787.
- 579 Kuang, H., Liu, J., Zeng, Y., Zhou, W., Wu, P., Tan, J., Li, Y., Pang, Q., Jiang, W., & Fan, R.
580 (2020). Co-exposure to polycyclic aromatic hydrocarbons, benzene and toluene may impair
581 lung function by increasing oxidative damage and airway inflammation in asthmatic children.

- 582 *Environmental Pollution*, 115220.
- 583 Kunst, A. E. (2021). Tobacco use in young people: being emic to end the epidemic. *The Lancet*
- 584 *Public Health*.
- 585 Lakier, J. B. (1992). Smoking and cardiovascular disease. *The American Journal of Medicine*,
- 586 93(1), S8–S12.
- 587 Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M.,
- 588 Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N.,
- 589 Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., Blore, J. D., ... Ezzati,
- 590 M. (2012). A comparative risk assessment of burden of disease and injury attributable to 67
- 591 risk factors and risk factor clusters in 21 regions, 1990–2010: A systematic analysis for the
- 592 Global Burden of Disease Study 2010. *The Lancet*. <https://doi.org/10.1016/S0140->
- 593 6736(12)61766-8
- 594 Maantay, J. (2002). Mapping environmental injustices: pitfalls and potential of geographic
- 595 information systems in assessing environmental health and equity. *Environmental Health*
- 596 *Perspectives*, 110(suppl 2), 161–171.
- 597 Martino-Andrade, A. J., & Chahoud, I. (2010). Reproductive toxicity of phthalate esters.
- 598 *Molecular Nutrition & Food Research*, 54(1), 148–157.
- 599 Mater, G., Paris, C., & Lavoué, J. (2016). Descriptive analysis and comparison of two French
- 600 occupational exposure databases: COLCHIC and SCOLA. *American Journal of Industrial*
- 601 *Medicine*, 59(5), 379–391.
- 602 Méndez-Armenta, M., & Ríos, C. (2007). Cadmium neurotoxicity. *Environmental Toxicology and*
- 603 *Pharmacology*, 23(3), 350–358.
- 604 Miao, H., & Ji, J. S. (2019). Trends of Blood Cadmium Concentration Among Workers and Non-

605 Workers in the United States (NHANES 2003 to 2012). *Journal of Occupational and*
606 *Environmental Medicine*, 61(12), e503–e509.

607 Minnesota Department of Health. (2020). *Volatile Organic Compounds in Your Home*.
608 <https://www.health.state.mn.us/communities/environment/air/toxins/voc.htm>

609 OECD. (2004). *The report from the expert group on (quantitative) structure-activity relationships*
610 *[(q) sars] on the principles for the validation of (q) sars*.

611 OECD. (2014). *Guidance document on the validation of (quantitative) structure-activity*
612 *relationship [(q)sar] models*.

613 Omenn, G. S., Goodman, G. E., Thornquist, M. D., Balmes, J., Cullen, M. R., Glass, A., Keogh,
614 J. P., Meyskens Jr, F. L., Valanis, B., & Williams Jr, J. H. (1996). Risk factors for lung cancer
615 and for intervention effects in CARET, the Beta-Carotene and Retinol Efficacy Trial. *JNCI:*
616 *Journal of the National Cancer Institute*, 88(21), 1550–1559.

617 Paolini, M., Cantelli-Forti, G., Perocco, P., Pedulli, G. F., Abdel-Rahman, S. Z., & Legator, M. S.
618 (1999). Co-carcinogenic effect of β -carotene. *Nature*, 398(6730), 760–761.

619 Program, N. T. (2006). Toxicology and carcinogenesis studies of benzophenone (CAS No. 119-
620 61-9) in F344/N rats and B6C3F1 mice (feed studies). *Natl. Toxicol. Program Tech. Rep. Ser*,
621 533, 1–264.

622 Raymond, C. V. J. (1998). Estimating the lung deposition of particulate polycyclic aromatic
623 hydrocarbons associated with multimodal urban aerosols. *Inhalation Toxicology*, 10(3), 183–
624 204.

625 Sadetzki, S., Chetrit, A., Turner, M. C., Van Tongeren, M., Benke, G., Figuerola, J., Fleming, S.,
626 Hours, M., Kincl, L., & Krewski, D. (2016). Occupational exposure to metals and risk of
627 meningioma: a multinational case-control study. *Journal of Neuro-Oncology*, 130(3), 505–

628 515.

629 Saetta, M., Di Stefano, A., Maestrelli, P., de Marzo, N., Milani, G. F., Pivrotto, F., Mapp, C. E.,
630 & Fabbri, L. M. (1992). Airway mucosal inflammation in occupational asthma induced by
631 toluene diisocyanate. *American Review of Respiratory Disease*, *145*(1), 160–168.

632 Schober, W., Lubitz, S., Belloni, B., Gebauer, G., Lintelmann, J., Matuschek, G., Weichenmeier,
633 I., Eberlein-König, B., Buters, J., & Behrendt, H. (2007). Environmental polycyclic aromatic
634 hydrocarbons (PAHs) enhance allergic inflammation by acting on human basophils.
635 *Inhalation Toxicology*, *19*(sup1), 151–156.

636 Shim, Y. K., Lewin, M. D., Ruiz, P., Eichner, J. E., & Mumtaz, M. M. (2017). Prevalence and
637 associated demographic characteristics of exposure to multiple metals and their species in
638 human populations: The United States NHANES, 2007–2012. *Journal of Toxicology and*
639 *Environmental Health, Part A*, *80*(9), 502–512.

640 Shimshack, J. P., Ward, M. B., & Beatty, T. K. M. (2007). Mercury advisories: information,
641 education, and fish consumption. *Journal of Environmental Economics and Management*,
642 *53*(2), 158–179.

643 Soley-Bori, M. (2013). Dealing with missing data: Key assumptions and methods for applied
644 analysis. *Boston University*, *4*, 1–19.

645 Stanaway, J. D., Afshin, A., Gakidou, E., Lim, S. S., Abate, D., Abate, K. H., Abbafati, C., Abbasi,
646 N., Abbastabar, H., & Abd-Allah, F. (2018). Global, regional, and national comparative risk
647 assessment of 84 behavioural, environmental and occupational, and metabolic risks or
648 clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the
649 Global Burden of Disease Stu. *The Lancet*, *392*(10159), 1923–1994.

650 Stillman, R. J., Rosenberg, M. J., & Sachs, B. P. (1986). Smoking and reproduction. *Fertility and*

- 651 *Sterility*, 46(4), 545–566.
- 652 Swartz, M. A. (2001). Occupation lead exposure: Health effects and remediation practices.
- 653 *Professional Safety*, 46(7), 28.
- 654 Trzcinka-Ochocka, M., Gazewski, A., & Brodzka, R. (2007). Exposure to mercury vapors in dental
- 655 workers in Poland. *International Journal of Occupational Medicine and Environmental*
- 656 *Health*, 20(2), 147–154.
- 657 Tyrrell, J., Melzer, D., Henley, W., Galloway, T. S., & Osborne, N. J. (2013). Associations
- 658 between socioeconomic status and environmental toxicant concentrations in adults in the
- 659 USA: NHANES 2001–2010. *Environment International*, 59, 328–335.
- 660 U.S. Department of Labor. (2019). *Fact Sheet #171: Blue-Collar Workers and the Part 541*
- 661 *Exemptions Under the Fair Labor Standards Act (FLSA)*.
- 662 Yin, S., Tang, M., Chen, F., Li, T., & Liu, W. (2017). Environmental exposure to polycyclic
- 663 aromatic hydrocarbons (PAHs): The correlation with and impact on reproductive hormones
- 664 in umbilical cord serum. *Environmental Pollution*, 220, 1429–1437.
- 665