

1 **Integrating exposure knowledge and serum suspect screening as a new approach to**
2 **biomonitoring: An application in firefighters and office workers**

3
4 Rachel Grashow¹, Vincent Bessonneau¹, Roy R. Gerona², Aolin Wang³, Jessica Trowbridge⁴,
5 Thomas Lin², Heather Buren⁶, Ruthann A. Rudel^{1*}, Rachel Morello-Frosch^{4,5*}

6 ¹Silent Spring Institute, Newton, MA, USA

7 ²Clinical Toxicology and Environmental Biomonitoring Lab, Department of Obstetrics,
8 Gynecology and Reproductive Sciences, University of California, San Francisco, California,
9 USA.

10 ³Program on Reproductive Health and the Environment, Department of Obstetrics, Gynecology
11 and Reproductive Sciences & Bakar Computational Health Sciences Institute, University of
12 California, San Francisco, California, USA

13 ⁴School of Public Health, University of California, Berkeley, California, USA

14 ⁵Department of Environmental Science, Policy and Management University of California,
15 Berkeley, California, USA

16 ⁶United Fire Service Women, San Francisco, California, USA

17

18 * Joint corresponding authors

19

20

21 Acknowledgements: The authors thank all of the WFBC study participants for their contribution
22 to the study. This work is supported by the California Breast Cancer Research Program #19BB-
23 2900 (RG,VB, RRG, JT, TL, HB, RAR, RMF), the National Institute of Environmental Health
24 Sciences R01ES027051 (AW, RMF) and the San Francisco Firefighter Cancer Prevention
25 Foundation (HB). We thank Anthony Stefani, Emily O'Rourke, Nancy Carmona, Karen Kerr,
26 Julie Mau, Natasha Parks, Lisa Holdcroft, SF Fire Chief Joanne Hayes-White, Incoming SF Fire
27 Chief Jeanine Nicholson, Sharyle Patton, Connie Engel and Nancy Buermeyer for their
28 contributions to the study.

29 RAR, RG, and VB, are employed at the Silent Spring Institute, a scientific research organization
30 dedicated to studying environmental factors in women's health. The Institute is a 501(c)3 public
31 charity funded by federal grants and contracts, foundation grants, and private donations,
32 including from breast cancer organizations. HB is former president and member of United Fire
33 Service Women, a 501(c)3 public charity dedicated to supporting the welfare of women in the
34 San Francisco Fire Department.

35

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

37 Corresponding author, email, and mailing address:

38 Jessica Trowbridge, jesstro@berkeley.edu, (510)289-5949

39 SHE lab – Rachel Morello-Frosch

40 130 Mulford, Hall #3144, University of California, Berkeley

41 Berkeley CA, 94720

42 **ABSTRACT**

43

44 **Background:** Women firefighters are exposed to recognized and probable carcinogens, yet there
45 are few studies of chemical exposures and associated health concerns, such as breast cancer.
46 Biomonitoring often requires *a priori* selection of compounds to be measured, and so may not
47 detect important, lesser known, exposures.

48 **Objectives:** The Women Firefighters Biomonitoring Collaborative (WFBC) created a biological
49 sample archive and conducted a general suspect screen (GSS) to address this data gap.

50 **Methods:** Using liquid chromatography-quadrupole time-of-flight mass spectrometry (LC-
51 QTOF/MS) we sought to identify candidate chemicals of interest in serum samples from 83
52 women firefighters (FF) and 79 office workers (OW) in San Francisco. Through the GSS
53 approach we identified chemical peaks by matching accurate mass from serum samples against a
54 custom chemical database of 740 slightly polar phenolic and acidic compounds, including many
55 of relevance to firefighting or breast cancer etiology. We then selected chemicals for
56 confirmation based on *a priori* criteria: 1) detection frequency or peak area differences between
57 OW and FF; 2) evidence of mammary carcinogenicity, estrogenicity, or genotoxicity; and 3) not
58 currently measured in large biomonitoring studies.

59 **Results:** We detected 620 chemicals that matched 300 molecular formulas in the WFBC
60 database, including phthalate metabolites, phosphate flame retardant metabolites, phenols,
61 pesticides, nitro- and nitroso-compounds, and per- and polyfluoroalkyl substances. The average
62 number of chemicals from the database that were detected in participants was 72 and 70 in FF
63 and OW, respectively. We confirmed 8 of the 20 prioritized suspect chemicals –including two
64 alkylphenols, ethyl paraben, BPF, PFOSAA, benzophenone-3, benzyl p-hydroxybenzoate, and
65 triphenyl phosphate--by running a matrix spike of the reference standards and using m/z,
66 retention time and the confirmation of at least two fragment ions as criteria for matching.

67 **Conclusion:** GSS provides a powerful high-throughput approach to identify and prioritize novel
68 chemicals for biomonitoring and health studies.

69

70 INTRODUCTION

71 Firefighters are exposed to complex and variable chemical mixtures that include known
72 carcinogens. In addition to exposures during fire suppression activities (Adetona et al. 2013; Fent
73 et al. 2014, 2018; Navarro et al. 2017; Pleil et al. 2014), firefighters pick up chemical exposures
74 from their equipment, such as fire extinguishing foams or protective gear (Alexander and Baxter
75 2016; Fent et al. 2015), and also from automotive diesel (Oliveira et al. 2017). These compounds
76 include benzene, polycyclic aromatic hydrocarbons (PAHs), nitro-PAHs, formaldehyde, dioxins,
77 flame retardants, polychlorinated biphenyls, and poly- and perfluorinated substances (PFAS)
78 (Caux et al. 2002; Feunekes et al. 1997; Moen and Ovrebø 1997; Waldman et al. 2016). These
79 chemicals are associated with a wide range of cancers and other health effects in human and
80 experimental animal studies, and it is noteworthy that many of these exposures have been
81 identified as potential breast carcinogens either because they cause mammary gland tumors in
82 laboratory animals, or because they alter mammary gland development (Rudel et al. 2011, 2014).

83 Research examining the chemical exposures and health risks faced by firefighters, and women
84 firefighters in particular, is limited. A 2015 study conducted by the National Institute for
85 Occupational Safety and Health (NIOSH) on 19,309 male US firefighters observed positive
86 associations between the total time spent at fires and lung cancer incidence and mortality, and
87 between the total number of response to fires and leukemia mortality from 1950-2009 (Daniels et
88 al. 2015). An earlier report from this NIOSH cohort that included 991 women showed non-
89 significant increases in breast cancer incidence and mortality in both men and women, compared
90 with the general US population; these increases were largest at younger ages (<65 for men, 50-55
91 for women) (Daniels et al. 2014). Studies in multiple countries have also documented an elevated
92 risk of certain cancers in male firefighters and other first responders, including thyroid, bladder,
93 kidney, prostate, testicular, breast, brain, digestive cancers, multiple myeloma, and non-
94 Hodgkin's lymphoma (Ahn et al. 2012; Bates 2007; Delahunt et al. 1995; Kang et al. 2008; Ma
95 et al. 2005, 2006; Tsai et al. 2015). A meta-analysis of 32 studies determined an increased risk of
96 certain cancers in the mostly male firefighter population (LeMasters et al. 2006). Most studies do
97 not calculate risks to female firefighters; however, in a study on cancer incidence among Florida
98 professional firefighters, female firefighters showed a significantly increased risk of cancer
99 overall, as well as Hodgkin's lymphoma disease and thyroid cancer, compared with the Florida
100 general population (Ma et al. 2006). Although women make up 5.1% of firefighters across the
101 United States, (US Department of Labor 2018) their numbers can be higher in urban
102 jurisdictions, including in San Francisco, which has one of the highest proportions of women
103 firefighters (15%) (Hulett et al. 2008). As fire departments diversify and increase the number of
104 women in their ranks, it is important to characterize chemical exposures and implications for
105 health outcomes of particular relevance to women, such as breast cancer, that might not be
106 addressed in existing studies, which have been primarily conducted among men.

107 Biomonitoring is an important tool in environmental and occupational health studies seeking to
108 link health outcomes to chemical exposures. External measurements including in air, dust, and
109 water do not always reflect internal dose, and biomonitoring studies in human tissue can
110 integrate over multiple routes of exposure including dermal, inhalation and ingestion. One

111 limitation of many biomonitoring studies is that they rely on *a priori* selection of targeted
112 chemicals for study. This *a priori* selection approach often lacks critical information about which
113 chemicals are present in occupational settings (Egeghy et al. 2012; Judson et al. 2009), and about
114 metabolic transformations. As a result, significant time and resources may be expended to
115 develop analytical methods to measure chemicals without knowing whether they are present in
116 biological specimens. For example, 20% of the 250 chemicals biomonitoring in NHANES since
117 1999 were not detected in 95% or more of the US population, indicating that the criteria for
118 selecting chemicals for biomonitoring has not always identified chemicals with prevalent
119 exposure (CDC 2009). A more efficient and systematic approach is needed to identify a broader
120 spectrum of environmental chemicals present in the human body; this strategy is now recognized
121 as a critical component of an “exposome” approach (Buck Louis et al. 2013; Rappaport 2011;
122 Wild 2012). One way to characterize the human exposome is to perform a general suspect screen
123 (GSS) of biospecimens using high-resolution mass spectrometry. Recent applications of this
124 approach identified novel chemical exposures among pregnant women, including benzophenone-
125 1 and bisphenol S (Gerona et al. 2018; Wang et al. 2018).

126 To better understand how women firefighters are exposed to potential breast carcinogens and
127 other understudied chemicals, we undertook a community-based, participatory biomonitoring
128 project, a partnership among firefighters, environmental health scientists and environmental
129 health advocates, known as the Women Firefighters Biomonitoring Collaborative, to develop a
130 biospecimen archive of women firefighters and office workers in San Francisco. As part of the
131 WFBC, we conducted a cross-sectional chemical biomonitoring study to identify novel chemical
132 exposures by applying a discovery-driven, general suspect screen (GSS) using high-resolution
133 mass spectrometry. Our goal was to characterize multiple chemical exposures, assess whether
134 these exposures differ between firefighters and office workers, and prioritize candidate
135 compounds for confirmation and targeted methods development. Ultimately, we applied a GSS
136 approach to advance discovery of novel environmental chemicals in human biomonitoring.

137 **METHODS**

138 *Study design*

139 The Women Firefighter Biomonitoring Collaborative (WFBC) was designed to measure and
140 compare exposures to potential breast carcinogens and other endocrine disrupting compounds
141 (EDCs) in two occupational cohorts--women firefighters (FF) and office workers (OW) from the
142 City of San Francisco, California, and to create an archive of biological specimens for
143 exposomics research. The GSS was performed on serum samples collected from female
144 firefighters and office workers using liquid chromatography-quadrupole time-of-flight mass
145 spectrometry (LC-QTOF/MS) to characterize a wide spectrum of exposures to candidate
146 compounds in our study population. This method screens for hundreds of acidic or phenolic
147 organic compounds of interest, so the results represent a significantly larger universe of
148 compounds in a biospecimen rather than a limited set of chemicals selected, *a priori*, for
149 quantification. Accurate mass of each unique molecule (i.e. mass-to-charge ratio, *m/z*) generated
150 by the LC-QTOF/MS was matched to chemical formulas from a custom database of 740
151 chemicals of interest, based on their relevance to firefighting and breast cancer etiology. From

152 this WFBC database, we compared detection frequencies and peak areas of candidate
153 compounds between firefighters and office workers to identify those that might be work-related.
154 We then systematically combined expert knowledge on the sources, uses and toxicity of
155 candidate compounds to prioritize and select a subset of chemicals for confirmation. Ultimately,
156 we sought to demonstrate how GSS methods can be used to improve efficiencies in human
157 biomonitoring by broadening the spectrum of potential environmental chemical exposures and
158 applying exposure science expertise to identify and prioritize specific chemicals for confirmation
159 by targeted analysis.

160 *Recruitment and consent*

161 Women were eligible to participate in the WFBC study if they were over 18 years old, non-
162 smokers, and employees of the City and County of San Francisco (office workers) or the San
163 Francisco Fire Department (firefighters). In addition, firefighters had to have been working
164 active duty for at least five years with the Department. Firefighters were recruited through letters,
165 emails, and phone calls that targeted firefighter organizations, including United Fire Service
166 Women, Local 798 of the International Association of Firefighters (IAFF), the Black Firefighters
167 Association, Asian Firefighters Association, and Los Bomberos (Latino Firefighter Association).
168 Informational meetings were held at the San Francisco Fire Department. Female office
169 employees with the City and County of San Francisco were recruited through informational
170 meetings, direct email, letters, telephone calls and by networking efforts through SEIU Local
171 1021. The study was publicized through regular newsletters and other online communication
172 outlets regularly sent to firefighters and other San Francisco City and County employees through
173 the Health Services System. WFBC study protocols were approved by the Institutional Review
174 Board of the University of California, Berkeley (Protocol # 2013-07-5512). Informed consent
175 was obtained prior to all interviews and sample collections. Subjects were not paid for
176 participation, but did receive a \$20.00 gift card and reimbursement to offset the cost of parking
177 and transportation. Blood samples were collected between June 2014 and March 2015.

178 *Interviews and sample collection*

179 Once consented and enrolled, participants were scheduled for an in-person interview and blood
180 collection. Subjects met with a member of the research team to answer questions about their diet,
181 home, job, other activities, and education. After completing the exposure interview, a trained
182 phlebotomist drew blood samples, which were collected in four 10 mL red-top tubes without
183 additives. Samples were collected at sites near participants' work site and transported in a cooler
184 with ice for processing within 3 hours of collection. Serum was separated by allowing it to clot at
185 room temperature, then centrifuging at 3000 rpm for 10 minutes and -4°C. Serum was aliquoted
186 into 1.2 mL cryo-vial tubes and stored at -80°C until analysis. All samples were processed and
187 analyzed at the University of California, San Francisco. We collected and processed samples
188 from 86 firefighters and 84 office workers. We analyzed serum samples from those who had
189 sufficient serum for the chemicals analysis: from 83 firefighters and 79 office worker
190 participants.

191 *WFBC suspect chemical database*

192 To build a chemical database for our general suspect screen, we began with a database of 696
193 chemicals developed previously to identify environmental organic acids (EOA) among pregnant
194 women, including chemicals from the following classes: phenols, such as parabens; phenolic and
195 acidic pesticides and their predicted acidic and phenolic metabolites; per- and polyfluoroalkyl
196 substances (PFAS); phthalate metabolites; phenolic metabolites of polybrominated diphenyl
197 ethers (OH-BDEs) and polychlorinated biphenyls (OH-PCBs) (Wang et al. 2018). These EOAs
198 include many common consumer product chemicals and environmental pollutants, as well as
199 356 predicted metabolites of common pesticides (Wang et al. 2018). We extended this EOA
200 database for our WFBC analysis by adding environmental chemicals that were relevant to
201 occupational exposures faced by firefighters and office workers and also chemicals implicated in
202 breast carcinogenesis based on toxicological evidence. Specifically, we assessed the viability of
203 adding over 100 chemicals, based on the following criteria: 1) chemicals shown to be rodent
204 mammary gland carcinogens or that affect mammary gland development and so may increase
205 breast cancer risk (Rudel et al. 2011, 2014); or 2) chemicals related to firefighting that could lead
206 to occupational exposures, including perfluorinated compounds found in firefighting foams, and
207 other flame retardants and their metabolites (Dodson et al. 2012, 2014; Rodgers et al. 2018).
208 Chemicals that fit these two criteria were added to the WFBC database if their structures were
209 expected to be compatible with the LC-QTOF/MS operating in negative ionization mode. For
210 example, carcinogenic PAHs were not added to the database because they are unlikely be
211 detected using this method. We were able to add 44 chemicals for a total of 740 in the WFBC
212 database (Table S1).

213 *General suspect screening analysis using liquid-chromatography and quadrupole time-of-flight* 214 *mass spectrometry (LC-QTOF/MS)*

215 General suspect screening of serum was performed as previously described (Gerona et al. 2018).
216 Briefly, 250 μ L of serum was spiked with 2.5 μ L of 1 mg/mL of internal standard (2.5 ng BPA-
217 d16) and centrifuged at 3,000 rpm for 10 min. Analytes were extracted using solid-phase
218 extraction (SPE; Waters Oasis HLB 10 mg, 1cc). Extracts were dried under a stream of nitrogen
219 gas and reconstituted in 250 μ L of 10% methanol.

220 Extracts were analyzed on a LC-QTOF/MS system consisting of an LC 1260 and a QTOF/MS
221 6550 (Agilent, Santa Cruz, CA, USA). Analytes were separated by reversed-phase
222 chromatography using a C18 column (Agilent Poroshell 120, 2.1 mm \times 100 mm, 2.7 mm particle
223 size) maintained at 55°C. Mobile phase A consisted of water with 0.05% ammonium acetate
224 (pH=7.8) and mobile phase B consisted of methanol with 0.05% ammonium acetate (pH=7.8).
225 The elution gradient employed was: 0-0.5 min, 5% B; 1.5 min, 30% B; 4.5 min, 70% B; 7.5-10
226 min, 100% B; 10.01-14 min, 5% B. The injection volume was 50 μ L.

227 Analyses were performed with a QTOF/MS operating in negative electrospray ionization mode
228 (ESI-). Ions were collected in the m/z 80–600 range at high resolution for eluates coming out of
229 the LC from 1-12 min. Using the Auto MS/MS mode (information-dependent acquisition), a
230 product ion scan (MS/MS) of the three most abundant peaks at high resolution was triggered
231 each time a precursor ion with an intensity of ≥ 500 counts/second was generated in the
232 QTOF/MS scan using a collision voltage ranging from 0 to 40 V depending on ions m/z . The LC-
233 QTOF/MS analysis produces a total ion chromatogram for each sample, which includes the
234 following: the accurate mass of each unique compound (expressed as m/z of their respective

235 anion), peak area, retention time (RT) and spectral data on the parent and fragment ions,
236 including isotopic pattern.

237 We used the Agilent MassHunter Qualitative Analysis software Find-by-formula (FBF)
238 algorithm to analyze QTOF/MS data for novel chemical exposures among firefighters and office
239 workers using a set of optimized parameters previously reported (Gerona et al. 2018). First, all
240 detected m/z were matched to potential compound hits in the WFBC chemical database. The
241 algorithm imports molecular formulas from the database, automatically calculates their m/z
242 values and then matches them to m/z measured by the QTOF/MS with a mass tolerance value of
243 10 ppm. A list of possible chemical matches was generated for all serum samples, which
244 included the accurate mass (m/z), mass error (i.e. the difference between the experimental and
245 the theoretical m/z), retention time (RT), peak area, and match scores (Schymanski et al. 2014).
246 The initial LC-QTOF full scan identification resulted in 12,051 features with unique retention
247 times, which matched to 300 chemical formulas in our WFBC database with multiple
248 RTs/formula, or 620 unique chemical formula/RT combinations.

249 *Retention time correction and isomer distinction*

250 Isomers (compounds with the same chemical formula but with different chemical structures) are
251 recognized by the LC-QTOF method as the presence of multiple RTs, (measured in minutes) per
252 chemical formula or mass. We distinguished isomers by clustering compounds based on RT.
253 Briefly, we first ranked all suspect detections by RT for each chemical formula. We considered a
254 suspect peak to be from a different isomer if its RT differed from the RT of the same chemical
255 formula in the previous row by more than 0.16 minutes. Cutoff points ranging from 0.15 to 0.20
256 with a 0.01 increment were tested, and 0.16 allowed the best distinction based on graphical
257 examination (Wang et al. 2018). Then, we aligned peaks originating from the same isomer to an
258 identical RT. The final analytical sample consisted of 4,791 suspect detections that matched to
259 620 suspect chemicals (i.e., unique combinations of chemical formula and retention time).

260 *Chemical selection for validation and confirmation*

261 We used a multi-step procedure and criteria to reduce the initial set of candidate chemical
262 matches from the LC-QTOF/MS to a smaller set of compounds for validation by prioritizing
263 matches that showed differences in exposure between firefighters and office workers or had
264 toxicity characteristics relevant to breast cancer. We focused our general suspect screen on
265 compounds in our database that were not pharmaceutical chemicals or chemicals that we had
266 already identified for targeted analysis. We then used the following initial criteria to prioritize
267 matches for validation: 1) at least 10% detection frequency difference between firefighters and
268 office workers; 2) a higher peak area (indicator of higher relative concentration) in firefighters
269 compared to office workers (paired t-test, $p \leq 0.1$); 3) ubiquitous chemicals detected in more than
270 90% of both firefighter and office worker groups and 4) whether a chemical had been flagged as
271 a mammary carcinogen or mammary gland developmental disruptor [in (Rudel et al. 2007,
272 2011)]. As shown in Figure 2, this process yielded an initial list of 71 chemicals that we then
273 narrowed down to 54 for potential confirmation based on the availability of an analytical
274 standard.

275 In a second step for prioritizing tentative chemical matches for validation, we scored the
276 remaining 54 chemicals based on the first set of selection criteria as well as the following
277 additional characteristics: flame retardant chemicals, chemicals identified as estrogenic or
278 genotoxic, chemicals not detected in office workers, and chemicals not currently biomonitoring in
279 NHANES (CDC 2019) or the California Biomonitoring Program (Biomonitoring California
280 2019) The specific criteria were chemicals: 1) listed as flame retardants [in (Dodson et al. 2012,
281 2014)]; 2) not detected in the office workers; 3) currently not biomonitoring in NHANES or
282 Biomonitoring California; 4) listed as “active” for at least one genotoxicity bioassay tested in
283 PubChem (Wang et al. 2017); 5) listed as “active” for at least one estrogen receptor bioassay in
284 PubChem (The PubChem Project). For bioassay data, results were downloaded from the
285 PubChem website for each chemical. Then assay descriptions were queried for terms including
286 “genotox*”, “estrogen” and “salmonella” (to flag all Ames assays). All assays matching those
287 terms listed as “active” were tallied and chemicals with active assays were prioritized.

288 We scored the chemicals by assigning one point for each of the nine criteria. The study team
289 reviewed the top scoring chemicals and selected twenty for validation based on score as well as
290 data on uses, toxicity and sources using the Comparative Toxicogenomics Database (CTDB)
291 (Davis et al. 2017), PubChem (Wang et al. 2017), Toxnet (Fowler and Schnall 2014), and the
292 Toxin and Toxin Target Database (T3DB) (Wishart et al. 2015) (Table S2). Peaks that matched
293 predicted pesticide metabolites in our database were not considered for validation because of the
294 additional uncertainty about their presence in biological samples and lack of available reference
295 standards.

296 *Confirmation of selected chemicals*

297 We confirmed the presence of suspect chemicals in the serum samples by running the LC-
298 QTOF/MS analysis using the corresponding reference standard spiked into synthetic serum.
299 Tentative chemical matches from participant samples were confirmed if the *m/z*, at least two
300 fragment peaks in the MS/MS spectra, and retention time of the authentic standard matched
301 those found in the serum samples, consistent with level 1 confidence in identification
302 (Schymanski et al. 2014).

303 *Statistical analysis*

304 For statistical comparisons across demographic and occupational groups, we used the Wilcoxon
305 rank sum test to compare continuous variables or the Fisher test for categorical variables. All
306 data analysis and visualizations were completed using R, version 3.3.2 (R Core Team 2015).

307 **RESULTS**

308 Table 1 shows the demographic characteristics for the 83 firefighters and the 79 office workers
309 recruited for the WFBC study. At the time of recruitment, the San Francisco Fire Department
310 (SFFD) had 224 active duty women firefighters who made up nearly 15% of its workforce.
311 Among our study population, the average age of women firefighters is 47.9 (± 4.6) years old and
312 the average time of service in the Department is 17.4 (± 4.2) years. The racial/ethnic make-up of
313 this population in the department is: 50% non-Hispanic White, 21% Asian/Pacific Islander, 17%
314 Hispanic/Latino, and 13% African American, which is reflected by recruited firefighter

315 participants. Among the office workers, the average age is 47 years old and most have worked an
316 average of 14.0 years for the City and County of San Francisco. The racial and ethnic make-up
317 of this workforce was statistically similar to that of the firefighters, with a higher percentage of
318 non-Hispanic Asian/Pacific Islanders (25%).

319 Overall, the firefighters and office workers were similar in terms of average age, race/ethnicity,
320 body mass index (BMI), parity, and hormone use. However, the household income for
321 firefighters was significantly higher when compared to office workers, probably because of the
322 relatively higher compensation rate for firefighting versus office or clerical work. There were
323 significantly more premenopausal women in the firefighter group. Finally, office workers had a
324 higher proportion of college graduates than the firefighters.

325

326 *Suspect screening analysis of serum samples*

327 Our general suspect screen analysis identified 12,051 candidate compounds across all serum
328 samples, which were then compared to 740 chemical formulas from the WFBC database.
329 Retention time correction identified 300 chemical formulas, with multiple retention times per
330 formula such that there were 620 putative chemicals in the firefighter and office worker samples.
331 These included phthalate metabolites, phosphate flame retardants (PFRs) and their metabolites,
332 phenols, pesticides, nitro- and nitroso- compounds, and per- and polyfluoroalkyl substances
333 (PFASs). Figure 1 shows the number of chemical suspect hits per participant for each chemical
334 class. A large number of chemicals detected in FF and OW using this analytical method were
335 phenols and phthalate metabolites. The average cumulative number of suspect chemicals
336 detected was 73 (minimum: 45, maximum: 109) and 70 (minimum: 45; maximum: 100) in FF
337 and OW, respectively. Thus, the non-targeted LC-QTOF/MS data acquisition in ESI- was able to
338 detect a wide range of suspect organic acids that include many common commercial chemicals.

339 *Chemical restriction and prioritization for validation*

340 We identified 71 chemicals that were: 1) more abundant in firefighters or 2) ubiquitous and not
341 already in NHANES or 3) tagged as a potential concern for breast cancer. Sixty-three of these
342 chemicals satisfied only one criteria, and eight satisfied more than one. We further reduced this
343 list to chemicals that had commercially available authentic standards, leaving 54 to be considered
344 for validation. These chemicals included phenols such as bisphenol F and some alkylphenols,
345 phthalate metabolites, PFAS, flame retardant metabolites, nitroso-compounds, and pesticides
346 (See Table S2). None of the chemicals had significantly different detection frequencies or peak
347 areas in FF versus OW, but many had smaller differences. Fewer than half were identified as
348 mammary carcinogens or developmental disruptors. We scored the remaining 54 chemicals
349 based on indications of toxicity and exposure potential (Figure 2, Table S2).

350 We selected chemicals for analytical validation after reviewing the priority scores across nine
351 criteria for the 54 chemicals along with data on uses, toxicity and sources (Table S2 provides this
352 information for all 71 candidate chemicals).

353 Table 2 shows the top 20 scoring candidate chemicals and indicates the priority rank and whether
354 the chemical was included in the confirmation testing. For example, 2,4-bis(1,1-dimethylethyl)
355 phenol had the top ranking, meeting six of the nine criteria (Table 2) and was selected for
356 validation. Three nitro- and nitroso compounds with high scores, including 1-ethylnitroso-3-(2-
357 oxopropyl)-urea, 1-ethylnitroso-3-(2-hydroxyethyl)-urea and 1-amyl-1- nitrosourea were
358 eliminated because although our initial search indicated standards were available, the cost to
359 purchase them was prohibitive. Bis(1,3-dichloro-2-propyl) phosphate (BDCIPP) was excluded
360 because it was already being targeted for analysis in this cohort. Estradiol was excluded because
361 it is endogenous and Nifurdazil, an anti-bacterial agent, was excluded because we were not
362 targeting pharmaceuticals. We included the remaining 14 priority chemicals in the confirmation
363 testing.

364

365 *Validation*

366 Authentic standards of the 14 selected chemicals were analyzed by LC-QTOF/MS to evaluate
367 their match with retention times and mass spectra in the samples. Retention times for chemical
368 candidates and authentic standards, exact masses, and validation status are listed in Table 3.
369 Eight chemicals were validated, including: 2,4-bis(1,1-dimethylethyl)phenol, 2-hydroxy-4-
370 methoxybenzophenone -2, bisphenol F, perfluorooctanesulfon-amidoacetate (PFOSAA),
371 diphenyl phosphate (DPP), ethyl-p-hydroxybenzoate (ethyl paraben), benzyl p-hydroxybenzoate
372 (PHBB), and 4-hexyloxyphenol.

373 We found that retention times in participants' serum did not match those of the standards for six
374 chemicals: 1-allyl-1-nitrosourea, 4-butoxyphenol, 2,3,6-trimethylphenol, 4-phenethylphenol, and
375 two isomers for 4-heptyloxyphenol.

376 **DISCUSSION**

377 The goal of this study was to apply a general suspect screening approach to identify novel
378 exposures to previously understudied chemicals – of particular relevance to firefighting and
379 breast cancer etiology -- among a cohort of women firefighters compared to office worker
380 controls. We used LC-QTOF/MS to screen for the presence of 740 chemicals of interest in serum
381 from women firefighters and office workers. Accurate masses of chemical suspects were
382 tentatively matched with exact masses from the WFBC chemical database developed for this
383 study; chemical suspects were then prioritized for validation based on criteria related to exposure
384 profiles between the two groups as well as toxicity information, expected exposure patterns, and
385 whether they are currently biomonitored in major surveillance programs or not.

386 We detected 620 chemicals that matched 300 different molecular formulas, including phthalate
387 metabolites, phosphate flame retardants and their metabolites, phenols, pesticides, nitro- and
388 nitroso-compounds, and PFAS in both FF and OW. The average number of suspect chemicals
389 detected was 73 and 70 in FF and OW, respectively. Eight of the 20 prioritized chemicals were
390 validated by analysis with a known standard and will ultimately be quantified in the samples.
391 This approach presents a novel and powerful method for using suspect screening in a cohort of
392 female firefighters to reveal exposures to previously unstudied chemicals and to prioritize
393 compounds for confirmation.

394 Among the eight chemicals whose identity was validated by matching retention time and MS/MS
395 fragmentation of a known standard, the results suggested that exposures were different between
396 firefighters and office workers for most of them, although the magnitude of the differences was
397 modest. Based on statistically significant differences in peak area, firefighters had higher relative
398 levels of exposure for 2,4-bis(1,1-dimethylethyl) phenol, and office workers for PFOSAA and
399 ethyl paraben (Table 2). Firefighters appeared to have slightly higher detection frequencies for
400 2-hydroxy-4-methoxybenzophenone (BP-3), bisphenol F, PFOSAA and ethyl paraben, and office
401 workers had a higher detection frequency for PHBB.

402 The validated chemicals included two phenols, (bisphenol F and PHBB), which are used as
403 bisphenol-A substitutes (Ng et al. 2015), and BP-3, which is a UV filter in sunscreens, textiles,
404 and other products. The chemical 2,4-bis(1,1-dimethylethyl) phenol (aka 2,4-di-tert butyl
405 phenol), is listed as a manufacturing chemical and a fuel additive, however since it was detected
406 in all of the participants it may have some common consumer use or be a metabolite of a

407 common exposure (CID 7311) (Kim et al. 2016). It is interesting to note the similarity to 4-tert
408 butyl phenol—a stronger estrogen mimic that is ubiquitous in residential settings (Rudel et al.
409 2003). Ethyl paraben is an antifungal preservative found in cosmetics, toys, sunscreen and
410 pesticides (Guo and Kannan 2013). A PFAS chemical, PFOSAA, was also validated. Previous
411 studies have reported higher firefighting exposures for PFASs (Laitinen et al. 2014; Rotander et
412 al. 2015), and findings of targeted analysis for PFASs in this cohort are forthcoming (Trowbridge
413 et al. in prep). Originally a metabolite of an active ingredient in Scotchgard stain and water
414 repellent, PFOSAA is listed as an automotive, construction-related and cleaning chemical, as
415 well as an inert pesticide ingredient (CID 23691014) (Kim et al. 2016). It may also be found in
416 firefighting foams. Diphenyl phosphate, a common metabolite of the flame retardant and
417 plasticizer triphenyl phosphate (Cooper et al. 2011), appeared to have similar concentrations in
418 firefighters and office workers.

419 Among the few studies previously conducted on firefighters, one (Waldman et al. 2016)
420 observed higher exposures to environmental phenols (i.e. bisphenol A, triclosan, benzophenone-
421 3 and methyl paraben) among Southern California firefighters compared to the general
422 population. Since this study also investigated firefighters from California, it is difficult to
423 decipher whether the prevalent exposures to phenols are specifically related to firefighting
424 activities or simply more prevalent among California populations in general.

425 The phenols and PFAS chemicals that were validated in this study have estrogenic activity
426 (Table 2) or are of concern for a diverse set of toxicity endpoints, such as effects on kidney,
427 liver, lipid metabolism, growth and development, mammary gland development, and
428 immunotoxicity (Post et al. 2017). While there were tentative matches to nitro and nitroso
429 chemicals, which are of interest because of their genotoxicity and carcinogenicity (Table 2), we
430 were not able to validate any of these compounds, either because the retention time did not
431 match the known standard or we could not obtain the standard.

432 The success of this general suspect screening technique to identify novel chemical exposures in
433 environmental and occupational health studies could be improved further if there were chemical
434 databases that contain mass spectral information about diverse chemicals of interest. Because
435 most public metabolomics databases, such as HMDB, Metlin or T3DB, contain few entries for
436 environmental chemicals (e.g. HMDB contains 163 entries for toxin/pollutant) and there are no
437 extensive mass spectral databases of environmental chemicals currently available, we instead
438 made comparisons to 740 chemicals in our database based on matching exact masses. This
439 approach allowed us to tentatively identify exposures of interest by focusing the search on a set
440 of chemicals of interest and for which the analytical method was optimized. We also
441 demonstrated that this approach can be effective in measuring low abundant chemicals in human
442 serum. For example, PFOS detected using the GSS (Table S2) was also confirmed and quantified
443 using targeted LC-MS/MS (median serum concentrations for the whole cohort were 4.1 ng/mL
444 for PFOS) (Trowbridge et al. in prep).

445 We were also interested in identifying exposures associated with work practices that are not
446 related to fire events, such as diesel fuel and exhaust from trucks and equipment in the station,
447 flame retardants and PFAS chemicals from firefighting foam and protective gear, chemicals used
448 to clean and gear, and possibly others. Some of the chemicals selected for targeted analyses may
449 be related to workplace exposures such as these, and this suspect screening approach is one way

450 to generate hypotheses about exposures and to prioritize novel compounds for confirmation and
451 quantification using targeted methods.

452 Our study has several limitations. The sample size is relatively modest, and a larger cohort would
453 have provided more power to detect candidate chemicals that differed between firefighters and
454 office workers. In addition, since most of chemicals we detected are non-persistent, we can
455 expect large intra-individual variability in serum due to temporal variation in exposure. Also,
456 only 15 firefighters had their blood sample collected within 24 hours of working at a fire event,
457 so it may be that the chemicals we detected were not necessarily associated with firefighting
458 activities. One way to better characterize chemicals originating from fighting fires would be to
459 perform a longitudinal analysis in which biospecimens would be collected before and after a fire
460 event (within 12-24h).

461 Our WFBC general suspect chemical database (740 chemicals) contained only a small fraction of
462 the chemicals that could be important exposures for firefighters and office workers and so we
463 may have missed some important compounds for this study population. The use of larger
464 chemical databases such as the EPA Distributed Structure-Searchable Toxicity (DSSTOX;
465 ~9,000 chemicals) (Richard and Williams 2002) or PubChem (~3,000 chemicals) (Kim et al.
466 2016) would provide detection of a larger set of chemical suspects. However, increasing the
467 number of chemicals in a general suspect database would likely also increase the number of
468 “hits” (tentative chemical RT matches), making it more challenging to confirm matches and
469 increasing the rate of false positives. Even with our database of 740 chemicals, six, two of which
470 are isomers, of the top 20 tentative chemical matches that we selected for validation showed a
471 retention time (RT) mismatch such that the study serum sample RT did not match the RT
472 generated from a reference standard. Combining LC-QTOF/MS data - collected using a data-
473 independent acquisition approach (i.e. MS/MS fragmentation of as many metabolites as possible
474 in a single acquisition) – with bioinformatics tools such as retention time prediction, in silico
475 MS/MS prediction and molecular networking analysis (Allard et al. 2016; Bessonneau et al.
476 2017) would help to address this issue. In addition, a careful validation of the chemical identity
477 using an authentic standard is required to avoid reporting false positive matches. Likewise, the
478 number of matching fragmentation peaks required to minimize false positives can be investigated
479 in future studies. Ultimately, the MS/MS spectra generated for any compound provide structural
480 information specific to a compound. This data becomes very valuable for distinguishing isomeric
481 compounds that may have very close retention times in chromatography.

482 Another limitation is that use of LC/QTOF-MS in negative ionization mode limited the types of
483 chemicals that could be detected to organic acids. The use of complementary platforms such as
484 LC-QTOF/MS in positive ionization mode or GC combined with high resolution MS would
485 expand the investigation to more diverse classes of chemicals. For example, Greer Wallace et al.
486 (Geer Wallace et al. 2017) identified several VOCs and PAHs in firefighters exposed to
487 controlled structure burns using targeted and non-targeted GC-MS analysis of exhaled breath
488 condensate. Some of these chemicals such as benzaldehyde and dimethyl sulfide have been
489 previously associated with smoke/fire and combustion sources while methyl tert-butyl ether is
490 commonly used as an additive to gasoline. Finally, some of the nitroso compounds with high
491 priority scores in our analysis such as 1-amyl-1-nitrosourea and 1-allyl-1-nitrosourea could not
492 be validated because standards were not available.

493 In summary, we present a general suspect screening approach based on LC-QTOF/MS that can
494 be used to identify novel chemical exposures (i.e. not previously biomonitored) in a way that is
495 not as strictly limited by *a priori* hypotheses required by targeted methods. The approach we
496 used to select chemicals for confirmation integrates information from the serum samples, toxicity
497 and usage databases and expert knowledge to direct attention to chemicals relevant to the health
498 of women firefighters, an understudied yet vulnerable occupational group. Follow-up studies
499 should include targeted analyses to confirm and quantify the identified chemicals in the cohort,
500 identification of potential sources of the exposures, extension of the approach to cover a broader
501 and more diverse chemical space, and assessment of potential associations with health outcomes
502 for validated chemicals.

503 **References**

- 504 Adetona O, Zhang JJ, Hall DB, Wang J-S, Vena JE, Naeher LP. 2013. Occupational exposure to
505 woodsmoke and oxidative stress in wildland firefighters. *Sci Total Environ* 449:269–275, PMID:
506 23434577, doi:10.1016/j.scitotenv.2013.01.075.
- 507 Ahn Y-S, Jeong K-S, Kim K-S. 2012. Cancer morbidity of professional emergency responders in Korea.
508 *Am J Ind Med* 55:768–778, PMID:22628010, doi:10.1002/ajim.22068.
- 509 Alexander BM, Baxter CS. 2016. Flame-retardant contamination of firefighter personal protective
510 clothing - A potential health risk for firefighters. *J Occup Environ Hyg* 13:D148-155,
511 PMID:27171467, doi:10.1080/15459624.2016.1183016.
- 512 Allard P-M, Péresse T, Bisson J, Gindro K, Marcourt L, Pham VC, et al. 2016. Integration of Molecular
513 Networking and In-Silico MS/MS Fragmentation for Natural Products Dereplication. *Anal Chem*
514 88:3317–3323, PMID: 26882108, doi:10.1021/acs.analchem.5b04804.
- 515 Bates MN. 2007. Registry-based case-control study of cancer in California firefighters. *Am J Ind Med*
516 50:339–344, PMID: 17427202, doi:10.1002/ajim.20446.
- 517 Bessonneau V, Ings J, McMaster M, Smith R, Bragg L, Servos M, et al. 2017. In vivo microsampling to
518 capture the elusive exposome. *Sci Rep* 7:44038, PMID: 28266605, doi:10.1038/srep44038.
- 519 Biomonitoring California. 2019. Results | Measuring Chemicals in Californians. Available:
520 [http://biomonitoring.ca.gov/results/chemical/all?field_chemical_name_target_id_selective\[0\]=16](http://biomonitoring.ca.gov/results/chemical/all?field_chemical_name_target_id_selective[0]=16)
521 1 [accessed 16 August 2016].
- 522 Buck Louis GM, Yeung E, Sundaram R, Laughon SK, Zhang C. 2013. The exposome--exciting
523 opportunities for discoveries in reproductive and perinatal epidemiology. *Paediatr Perinat*
524 *Epidemiol* 27:229–236, PMID: 23574410, doi:10.1111/ppe.12040.
- 525 Caux C, O'Brien C, Viau C. 2002. Determination of firefighter exposure to polycyclic aromatic
526 hydrocarbons and benzene during fire fighting using measurement of biological indicators. *Appl*
527 *Occup Environ Hyg* 17:379–386, PMID: 12018402, doi:10.1080/10473220252864987.
- 528 CDC. 2009. Fourth national report on human exposure to environmental chemicals.
529 <http://www.cdc.gov/ExposureReport/pdf/FourthReport.pdf>.
- 530 CDC. 2019. National Report on Human Exposure to Environmental Chemicals | CDC. *Cent Dis Control*
531 *Prev*. Available: <https://www.cdc.gov/exposurereport/index.html> [accessed 18 November 2018].
- 532 Cooper EM, Covaci A, van Nuijs ALN, Webster TF, Stapleton HM. 2011. Analysis of the flame retardant
533 metabolites bis(1,3-dichloro-2-propyl) phosphate (BDCPP) and diphenyl phosphate (DPP) in
534 urine using liquid chromatography-tandem mass spectrometry. *Anal Bioanal Chem* 401:2123–
535 2132, PMID: 21830137, doi:10.1007/s00216-011-5294-7.
- 536 Daniels RD, Bertke S, Dahm MM, Yiin JH, Kubale TL, Hales TR, et al. 2015. Exposure-response
537 relationships for select cancer and non-cancer health outcomes in a cohort of U.S. firefighters
538 from San Francisco, Chicago and Philadelphia (1950-2009). *Occup Environ Med* 72:699–706,
539 PMID: 25673342, doi:10.1136/oemed-2014-102671.

- 540 Daniels RD, Kubale TL, Yiin JH, Dahm MM, Hales TR, Baris D, et al. 2014. Mortality and cancer
541 incidence in a pooled cohort of US firefighters from San Francisco, Chicago and Philadelphia
542 (1950-2009). *Occup Environ Med* 71:388–397, PMID: 24142974, doi:10.1136/oemed-2013-
543 101662.
- 544 Davis AP, Grondin CJ, Johnson RJ, Sciaky D, King BL, McMorrin R, et al. 2017. The Comparative
545 Toxicogenomics Database: update 2017. *Nucleic Acids Res* 45:D972–D978, PMID: 27651457,
546 doi:10.1093/nar/gkw838.
- 547 Delahunt B, Bethwaite PB, Nacey JN. 1995. Occupational risk for renal cell carcinoma. A case-control
548 study based on the New Zealand Cancer Registry. *Br J Urol* 75: 578–582, PMID: 7613791, doi:
549 10.1111/j.1464-410X.1995.tb07410.x.
- 550 Dodson RE, Perovich LJ, Covaci A, Van den Eede N, Ionas AC, Dirtu AC, et al. 2012. After the PBDE
551 phase-out: a broad suite of flame retardants in repeat house dust samples from California. *Environ*
552 *Sci Technol* 46:13056–13066, PMID: 23185960, doi:10.1021/es303879n.
- 553 Dodson RE, Van den Eede N, Covaci A, Perovich LJ, Brody JG, Rudel RA. 2014. Urinary biomonitoring
554 of phosphate flame retardants: levels in California adults and recommendations for future studies.
555 *Environ Sci Technol* 48:13625–13633, PMID: 245388620, doi:10.1021/es503445c.
- 556 Egeghy PP, Judson R, Gangwal S, Mosher S, Smith D, Vail J, et al. 2012. The exposure data landscape
557 for manufactured chemicals. *Sci Total Environ* 414:159–166, PMID: 22104386,
558 doi:10.1016/j.scitotenv.2011.10.046.
- 559 Fent KW, Eisenberg J, Snawder J, Sammons D, Pleil JD, Stiegel MA, et al. 2014. Systemic exposure to
560 PAHs and benzene in firefighters suppressing controlled structure fires. *Ann Occup Hyg* 58:830–
561 845, PMID: 24906357, doi:10.1093/annhyg/meu036.
- 562 Fent KW, Evans DE, Babik K, Striley C, Bertke S, Kerber S, et al. 2018. Airborne contaminants during
563 controlled residential fires. *J Occup Environ Hyg* 15:399–412, PMID: 29494297,
564 doi:10.1080/15459624.2018.1445260.
- 565 Fent KW, Evans DE, Booher D, Pleil JD, Stiegel MA, Horn GP, et al. 2015. Volatile Organic Compounds
566 Off-gassing from Firefighters' Personal Protective Equipment Ensembles after Use. *J Occup*
567 *Environ Hyg* 12:404–414, PMID: 25751596, doi:10.1080/15459624.2015.1025135.
- 568 Feunekes FD, Jongeneelen FJ, vd Laan H, Schoonhof FH. 1997. Uptake of polycyclic aromatic
569 hydrocarbons among trainers in a fire-fighting training facility. *Am Ind Hyg Assoc J* 58:23–28,
570 PMID: 9018833, doi:10.1080/15428119791013035.
- 571 Fowler S, Schnall JG. 2014. TOXNET: information on toxicology and environmental health. *Am J Nurs*
572 114:61–63, PMID: 24481372, doi:10.1097/01.NAJ.0000443783.75162.79.
- 573 Geer Wallace MA, Pleil JD, Mentese S, Oliver KD, Whitaker DA, Fent KW. 2017. Calibration and
574 performance of synchronous SIM/scan mode for simultaneous targeted and discovery (non-
575 targeted) analysis of exhaled breath samples from firefighters. *J Chromatogr A* 1516:114–124,
576 PMID: 28838652, doi:10.1016/j.chroma.2017.07.082.
- 577 Gerona RR, Schwartz JM, Pan J, Friesen MM, Lin T, Woodruff TJ. 2018. Suspect screening of maternal
578 serum to identify new environmental chemical biomonitoring targets using liquid

- 579 chromatography-quadrupole time-of-flight mass spectrometry. *J Expo Sci Environ Epidemiol*
580 28:101–108, PMID: 29019345, doi:10.1038/jes.2017.28.
- 581 Guo Y, Kannan K. 2013. A survey of phthalates and parabens in personal care products from the United
582 States and its implications for human exposure. *Environ Sci Technol* 47:14442–14449, PMID:
583 24261694, doi:10.1021/es4042034.
- 584 Hulett DM, Bendick Jr. M, Thomas SY, Moccio F. 2008. Enhancing Women’s Inclusion in Firefighting
585 in the USA. *Int J Divers Organ Communities Nations* 8.
- 586 Judson R, Richard A, Dix DJ, Houck K, Martin M, Kavlock R, et al. 2009. The toxicity data landscape
587 for environmental chemicals. *Environ Health Perspect* 117:685–695, PMID: 19479008,
588 doi:10.1289/ehp.0800168.
- 589 Kang D, Davis LK, Hunt P, Kriebel D. 2008. Cancer incidence among male Massachusetts firefighters,
590 1987–2003. *Am J Ind Med* 51:329–335, PMID: 18306327, doi:10.1002/ajim.20549.
- 591 Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, et al. 2016. PubChem Substance and
592 Compound databases. *Nucleic Acids Res* 44:D1202–1213, PMID: 26400175,
593 doi:10.1093/nar/gkv951.
- 594 Laitinen JA, Koponen J, Koikkalainen J, Kiviranta H. 2014. Firefighters’ exposure to perfluoroalkyl acids
595 and 2-butoxyethanol present in firefighting foams. *Toxicol Lett* 231:227–232, PMID: 25447453,
596 doi:10.1016/j.toxlet.2014.09.007.
- 597 LeMasters GK, Genaidy AM, Succop P, Deddens J, Sobeih T, Barriera-Viruet H, et al. 2006. Cancer risk
598 among firefighters: a review and meta-analysis of 32 studies. *J Occup Environ Med* 48:1189–
599 1202, PMID: 17099456, doi:10.1097/01.jom.0000246229.68697.90.
- 600 Ma F, Fleming LE, Lee DJ, Trapido E, Gerace TA. 2006. Cancer incidence in Florida professional
601 firefighters, 1981 to 1999. *J Occup Environ Med* 48:883–888, PMID: 15898094,
602 doi:10.1097/01.jom.0000235862.12518.04.
- 603 Ma F, Fleming LE, Lee DJ, Trapido E, Gerace TA, Lai H, et al. 2005. Mortality in Florida professional
604 firefighters, 1972 to 1999. *Am J Ind Med* 47:509–517, PMID: 16966954,
605 doi:10.1002/ajim.20160.
- 606 Moen BE, Ovrebø S. 1997. Assessment of exposure to polycyclic aromatic hydrocarbons during
607 firefighting by measurement of urinary 1-hydroxypyrene. *J Occup Environ Med* 39: 515–519,
608 PMID: 9211208, doi: 10.1097/00043764-199706000-00005.
- 609 Navarro KM, Cisneros R, Noth EM, Balmes JR, Hammond SK. 2017. Occupational Exposure to
610 Polycyclic Aromatic Hydrocarbon of Wildland Firefighters at Prescribed and Wildland Fires.
611 *Environ Sci Technol* 51:6461–6469, PMID: 28498656, doi:10.1021/acs.est.7b00950.
- 612 Ng HW, Shu M, Luo H, Ye H, Ge W, Perkins R, et al. 2015. Estrogenic activity data extraction and in
613 silico prediction show the endocrine disruption potential of bisphenol A replacement compounds.
614 *Chem Res Toxicol* 28:1784–1795; PMID: 26308263, doi:10.1021/acs.chemrestox.5b00243.
- 615 Oliveira M, Slezakova K, Fernandes A, Teixeira JP, Delerue-Matos C, Pereira M do C, et al. 2017.
616 Occupational exposure of firefighters to polycyclic aromatic hydrocarbons in non-fire work

- 617 environments. *Sci Total Environ* 592:277–287, PMID: 28319714,
618 doi:10.1016/j.scitotenv.2017.03.081.
- 619 Pleil JD, Stiegel MA, Fent KW. 2014. Exploratory breath analyses for assessing toxic dermal exposures
620 of firefighters during suppression of structural burns. *J Breath Res* 8:037107, PMID: 25190461,
621 doi:10.1088/1752-7155/8/3/037107.
- 622 Post GB, Gleason JA, Cooper KR. 2017. Key scientific issues in developing drinking water guidelines for
623 perfluoroalkyl acids: Contaminants of emerging concern. *PLoS Biol* 15:e2002855, PMID:
624 29261653, doi:10.1371/journal.pbio.2002855.
- 625 R Core Team. 2015. *R: A language and environment for statistical computing*. R Foundation for
626 Statistical Computing:Vienna, Austria.
- 627 Rappaport SM. 2011. Implications of the exposome for exposure science. *J Expo Sci Environ Epidemiol*
628 21:5–9, PMID: 21081972, doi:10.1038/jes.2010.50.
- 629 Richard AM, Williams CR. 2002. Distributed structure-searchable toxicity (DSSTox) public database
630 network: a proposal. *Mutat Res* 499: 27–52, PMID: 11804603, doi:10.1016/S0027-
631 5107(01)00289-5
- 632 Rodgers KM, Udesky JO, Rudel RA, Brody JG. 2018. Environmental chemicals and breast cancer: An
633 updated review of epidemiological literature informed by biological mechanisms. *Environ Res*
634 160:152–182, PMID: 28987728, doi:10.1016/j.envres.2017.08.045.
- 635 Rotander A, Kärman A, Toms L-ML, Kay M, Mueller JF, Gómez Ramos MJ. 2015. Novel fluorinated
636 surfactants tentatively identified in firefighters using liquid chromatography quadrupole time-of-
637 flight tandem mass spectrometry and a case-control approach. *Environ Sci Technol* 49:2434–
638 2442, PMID: 25611076, doi:10.1021/es503653n.
- 639 Rudel RA, Ackerman JM, Attfield KR, Brody JG. 2014. New exposure biomarkers as tools for breast
640 cancer epidemiology, biomonitoring, and prevention: a systematic approach based on animal
641 evidence. *Environ Health Perspect* 122:881–895, PMID: 24818537, doi:10.1289/ehp.1307455.
- 642 Rudel RA, Attfield KR, Schifano JN, Brody JG. 2007. Chemicals causing mammary gland tumors in
643 animals signal new directions for epidemiology, chemicals testing, and risk assessment for breast
644 cancer prevention. *Cancer* 109:2635–2666, PMID: 17503434, doi:10.1002/cncr.22653.
- 645 Rudel RA, Camann DE, Spengler JD, Korn LR, Brody JG. 2003. Phthalates, alkylphenols, pesticides,
646 polybrominated diphenyl ethers, and other endocrine-disrupting compounds in indoor air and
647 dust. *Environ Sci Technol* 37: 4543–4553, PMID: 14594359, doi: 10.1021/es0264596.
- 648 Rudel RA, Fenton SE, Ackerman JM, Euling SY, Makris SL. 2011. Environmental exposures and
649 mammary gland development: state of the science, public health implications, and research
650 recommendations. *Environ Health Perspect* 119:1053–1061, PMID: 21697028,
651 doi:10.1289/ehp.1002864.
- 652 Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, et al. 2014. Identifying Small Molecules
653 via High Resolution Mass Spectrometry: Communicating Confidence. *Env Sci Technol* 48:2097–
654 2098, PMID: 24476540, doi:10.1021/es5002105.

- 655 The PubChem Project. Available: <https://pubchem.ncbi.nlm.nih.gov/> [accessed 19 March 2019].
- 656 Trowbridge J, Gerona RR, Lin T, Rudel RA, Bessonneau V, Buren H, et al. in prep. Exposure to
657 Perfluoroalkyl substances in a cohort of women firefighters and office workers in San Francisco.
658 In Prep.
- 659 Tsai RJ, Luckhaupt SE, Schumacher P, Cress RD, Deapen DM, Calvert GM. 2015. Risk of cancer among
660 firefighters in California, 1988-2007. *Am J Ind Med* 58:715–729, PMID: 25943908,
661 doi:10.1002/ajim.22466.
- 662 US Department of Labor B of LS. 2018. Employed persons by detailed occupation, sex, race, and
663 Hispanic or Latino ethnicity. *Labor Force Stat Curr Popul Surv* 2018. Available:
664 <https://www.bls.gov/cps/cpsaat11.htm> [accessed 19 March 2019].
- 665 Waldman JM, Gavin Q, Anderson M, Hoover S, Alvaran J, Ip HSS, et al. 2016. Exposures to
666 environmental phenols in Southern California firefighters and findings of elevated urinary
667 benzophenone-3 levels. *Environ Int* 88:281–287, PMID: 26821331,
668 doi:10.1016/j.envint.2015.11.014.
- 669 Wang A, Gerona RR, Schwartz JM, Lin T, Sirota M, Morello-Frosch R, et al. 2018. A Suspect Screening
670 Method for Characterizing Multiple Chemical Exposures among a Demographically Diverse
671 Population of Pregnant Women in San Francisco. *Environ Health Perspect* 126:077009, PMID:
672 30044231, doi:10.1289/EHP2920.
- 673 Wang Y, Bryant SH, Cheng T, Wang J, Gindulyte A, Shoemaker BA, et al. 2017. PubChem BioAssay:
674 2017 update. *Nucleic Acids Res* 45:D955–D963, PMID: 27899599, doi:10.1093/nar/gkw1118.
- 675 Wild CP. 2012. The exposome: from concept to utility. *Int J Epidemiol* 41:24–32, PMID: 22296988,
676 doi:10.1093/ije/dyr236.
- 677 Wishart D, Arndt D, Pon A, Sajed T, Guo AC, Djoumbou Y, et al. 2015. T3DB: the toxic exposome
678 database. *Nucleic Acids Res* 43:D928-934, PMID: 25378312, doi:10.1093/nar/gku1004.
- 679

680 **Table 1. WFBC study population characteristics**

Characteristic	Office Workers (n=79)	Firefighters (n=83)	p-value^a
Age			
Mean ± SD	48.1 ± 10.6	47.9 ± 8.4	0.4
Race/ethnicity <i>n</i> (%)			
Non-Hispanic Asian	17 (22)	13 (16)	0.3
Non-Hispanic blacks	5 (6)	9 (11)	
Hispanics of all races	7 (9)	8 (9)	
Multiracial	10 (13)	16 (19)	
Non-Hispanic whites	40 (50)	37 (45)	
Education <i>n</i> (%)			
High school or less	5 (6)	6 (7)	< 0.001
Some college	10 (13)	40 (48)	
College graduates or higher	64 (81)	37 (45)	
BMI			
Mean (SD)	25.8 (5.2)	26.2 (3.5)	0.2
Household income <i>n</i> (%)			
< \$99,999	23 (29)	1 (1)	< 0.001
\$100,000-174,999	18 (23)	29 (35)	
\$175,000-199,999	12 (15)	17 (20)	
> \$200,000	26 (33)	36 (44)	
Menopausal status <i>n</i> (%)			
Premenopausal	44 (56)	62 (75)	0.007
Postmenopausal	35 (44)	21 (25)	
Hormone use ^b <i>n</i> (%)			
Never	19 (26)	16 (20)	0.6
During the past	38 (53)	46 (60)	

Currently	15 (21)	15 (20)	
Parity (# of live births) <i>n</i> (%)			
0	36 (46)	34 (41)	0.3
1	18 (23)	15 (18)	
> 1	25 (31)	34 (41)	

681 SD: Standard deviation; ^a Wilcoxon rank sum test to compare continuous variables by firefighter
682 status or Fisher test for categorical variables; ^b Missing data on hormone use for 6 firefighters
683 and 7 office workers.

Table 2: Twenty highest scoring chemicals prioritized for validation

Chemical name	Class	Rank	DF FF (%)	DF OW (%)	Mean peak area FF	Mean peak area OW	Flame retardant	DF > 90% in FF and OW	DF_FF - DF_OW > 10%	T-test PA p<0.1	Unmonitored ^a	Genotoxic	Estrogenic	OW non-defect	MC list	Total Score	Validation status
2,4-bis(1,1-dimethylethyl)phenol	Phenol	1	82 (100%)	76 (100%)	9.17E+05 [†]	7.66E+05	0	1	0	1	1	1	1	0	0	5	S
benzyl p-hydroxybenzoate (PHBB) or ^b	Phenol	2	16 (19.5%)	6 (7.9%)	2.98E+04	2.12E+04	0	0	1	0	1	1	0	0	0		S
2-hydroxy-4-methoxybenzophenone -2 (BP-3)																	
4-hexyloxyphenol	Phenol	3	81 (98.8%)	71 (93.4%)	1.04E+05*	7.51E+04	0	1	0	1	1	1	1	0	0	5	S
benzyl p-hydroxybenzoate (PHBB) or ^b	Phenol	4	30 (36.6%)	38 (50%)	6.04E+04	9.68E+04	0	0	1	0	1	1	0	0	0		S
2-hydroxy-4-methoxybenzophenone -2																	
bisphenol F	Phenol	5	10 (12.2%)	0 (0%)	4.98E+05	NA	0	0	1	0	1	0	1	1	0	4	S
4-butoxyphenol	Phenol	6	77 (93.9%)	71 (93.4%)	7.21E+04	8.58E+04 [†]	0	1	0	1	1	0	0	0	0	3	S
2,3,6-trimethylphenol	Phenol	7	18 (22%)	7 (9.2%)	2.04E+04	1.15E+04	0	0	1	0	1	0	0	0	0	2	S
1-ethylnitroso-3-(2-oxopropyl)-urea	Nitro and Nitroso	8	14 (17.1%)	10 (13.2%)	2.54E+04	2.09E+04	0	0	0	0	1	0	0	0	1	2	E-No std

	Compound																
perfluorooctanesulfona midoacetate (PFOSAA)	PFAS	9	16 (19.5%)	25 (32.9%)	3.94E+0 4	4.56E+0 4†	0	0	1	1	1	0	0	0	0	3	S
diphenyl phosphate (DPP)	Phosphate Flame Retardant metabolite	10	45 (54.9%)	39 (51.3%)	1.57E+0 4	1.68E+0 4	1	0	0	0	0	0	0	0	0	1	S
bis(1,3-dichloro-2-propyl) phosphate (BDCIPP)	Phosphate Flame Retardant metabolite	11	2 (2.4%)	1 (1.3%)	1.35E+0 4	1.13E+0 4	1	0	0	0	0	0	0	0	1	2	E- target analyte
4-phenethylphenol	Phenol	12	82 (100%)	76 (100%)	1.35E+0 5	1.43E+0 5†	0	1	0	1	1	0	1	0	0	4	S
4-heptyloxyphenol ^b (isomer 1)	Phenol	13	31 (37.8%)	21 (27.6%)	6.60E+0 4	6.87E+0 4	0	0	1	0	1	0	1	0	0	3	S
Nifurdazil	Nitro and Nitroso Compound	14	4 (4.9%)	3 (3.9%)	2.37E+0 4	1.07E+0 4	0	0	0	0	1	1	0	0	1	3	E - medication
4-heptyloxyphenol ^b (isomer 2)	Phenol	15	51 (62.2%)	55 (72.4%)	2.89E+0 5	2.55E+0 5	0	0	1	0	1	0	1	0	0	3	S
1-ethylnitroso-3-(2-hydroxyethyl)-urea	Nitro and Nitroso Compound	16	3 (3.7%)	2 (2.6%)	1.57E+0 4	1.57E+0 4	0	0	0	0	1	0	0	0	1	2	E-No std
1-amyl-1- nitrosourea	Nitro and Nitroso Compound	17	7 (8.5%)	11 (14.5%)	3.56E+0 4	2.33E+0 4	0	0	0	0	1	0	0	0	1	2	E-No std
ethyl-p-hydroxybenzoate (ethyl paraben)	Phenol	18	52 (63.4%)	35 (46.1%)	1.10E+0 5	1.57E+0 5*	0	0	1	1	0	0	1	0	0	3	S
1-allyl-1-nitrosourea	Nitro and Nitroso Compound	19	12 (14.6%)	5 (6.6%)	7.25E+0 4	3.96E+0 4	0	0	0	0	1	0	0	0	1	2	S
estradiol	Steroid	20	1 (1.2%)	0 (0%)	1.03E+0 4	NA	0	0	0	0	0	1	1	1	1	4	E- endogenous

^a Unmonitored in NHANES or Biomonitoring California; ^b these are isomers and could not be distinguished based on molecular mass; *p<0.1; † p<0.05; FF = firefighter; OW = office worker; DF = detection frequency; PA = peak area; RT=retention time; MC=mammary carcinogen; E = eliminated for validation; S = selected for validation; LOD = limit of detection; std=standard

Table 3: Retention time and exact mass for chemicals selected for validation

Chemical name	Chemical class	# of isomers	Mean RT for serum samples	RT lab standard	Validation status
2,4-bis(1,1-dimethylethyl) phenol	Phenol	4	4.33, 5.25, 5.48, 6.73	6.72	✓
2-hydroxy-4-methoxybenzophenone (BP-3))	Phenol	2	4.33, 5.25	5.30	✓
bisphenol F	Phenol	2	3.91	4.00	✓
perfluorooctanesulfonamidoacetate (PFOSAA)	PFC	1	5.93	5.95	✓
diphenyl phosphate (DPP)	Phosphate Flame Retardant metabolite	1	3.86	3.90	✓
ethyl-p-hydroxybenzoate (ethyl paraben)	Phenol	2	2.21, 3.80	2.30	✓
benzyl p-hydroxybenzoate (PHBB)	Phenol	2	4.33, 5.25	4.40	✓
4-hexyloxyphenol ¹	Phenol	1	5.81	5.80	✓ ^a
4-butoxyphenol	Phenol	1	4.19	5.10	✗ ^b
2,3,6-trimethylphenol	Phenol	2	3.97	4.25	✗ ^b
4-phenethylphenol	Phenol	1	5.71	6.02	✗ ^b
4-heptyloxyphenol (2 isomers)	Phenol	1	5.09	6.22	✗ ^b
1-allyl-1-nitrosourea	Nitro and Nitroso compound	1	0.76	1.20	✗ ^b

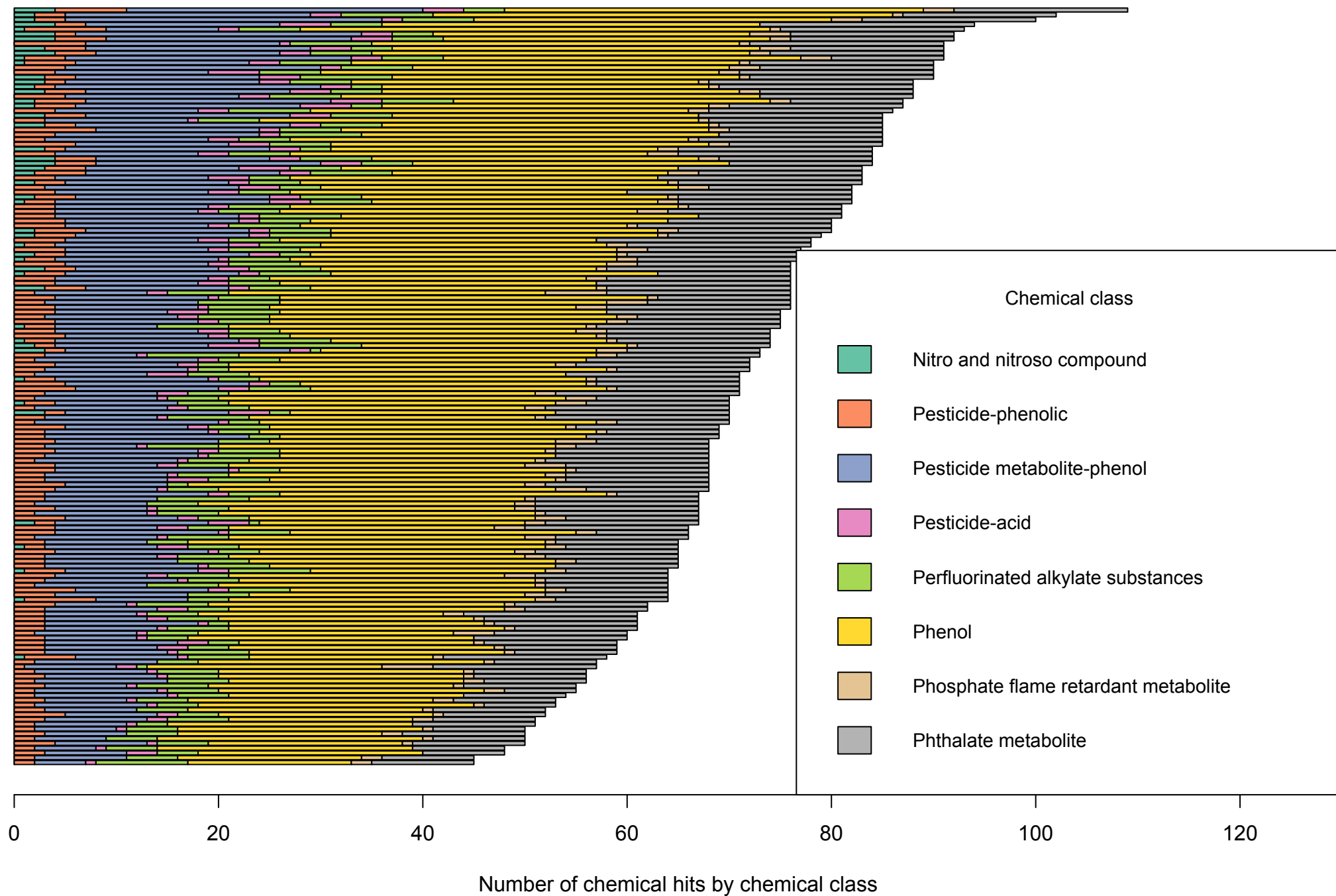
^a validated but with high LOD, ^b not validated because of retention time mismatch

Figure 1: Cumulative number of WFBC database chemicals detected with LC-QTOF/MS ESI- in serum samples from 162 study participants (mean=72; min=45; max=109).

Figure 2: Scoring and ranking of chemicals detected by LC-QTOF.

Figure 2 legend: PA= peak area; FF= firefighters; OW = office workers; DF = detection frequency, MC= mammary carcinogen; MGDD = mammary gland developmental disruptor

Participant ID



MS-library matching

Full scan of 162 samples yields 12,051 isomers with unique retention times

In-house library containing 740 chemical formulas

12,051 isomers matched to 300 chemical formulas (multiple retention times/formula)

Retention time correction identifies 620 putative chemicals

Candidate restriction

Inclusion criteria

↑ Peak area in FF	DF difference	Ubiquitous detection	Chemical of interest
$PA_{FF} > PA_{OW}$ $p < 0.1$ (n = 7)	$ DF_{FF} - DF_{OW} > 10\%$ (n = 36)	$DF_{OW} > 90\% \cap DF_{FF} > 90\%$ (n = 13)	MC or MGDD (n = 23)

71 candidate chemicals for validation

54 commercially available

Scoring criteria

