



Non-targeted GC/MS analysis of exhaled breath samples: Exploring human biomarkers of exogenous exposure and endogenous response from professional firefighting activity

M. Ariel Geer Wallace, Joachim D. Pleil, Karen D. Oliver, Donald A. Whitaker, Sibel Mentese, Kenneth W. Fent & Gavin P. Horn

To cite this article: M. Ariel Geer Wallace, Joachim D. Pleil, Karen D. Oliver, Donald A. Whitaker, Sibel Mentese, Kenneth W. Fent & Gavin P. Horn (2019) Non-targeted GC/MS analysis of exhaled breath samples: Exploring human biomarkers of exogenous exposure and endogenous response from professional firefighting activity, *Journal of Toxicology and Environmental Health, Part A*, 82:4, 244-260, DOI: [10.1080/15287394.2019.1587901](https://doi.org/10.1080/15287394.2019.1587901)

To link to this article: <https://doi.org/10.1080/15287394.2019.1587901>



View supplementary material 



Published online: 23 Mar 2019.



Submit your article to this journal 



Article views: 15



View Crossmark data 



Non-targeted GC/MS analysis of exhaled breath samples: Exploring human biomarkers of exogenous exposure and endogenous response from professional firefighting activity

M. Ariel Geer Wallace^a, Joachim D. Pleil^a, Karen D. Oliver^a, Donald A. Whitaker^a, Sibel Mentese^b, Kenneth W. Fent^c, and Gavin P. Horn^d

^aOffice of Research and Development, National Exposure Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC, USA; ^bDepartment of Environmental Engineering, Çanakkale Onsekiz Mart University, Merkez/Çanakkale, Turkey; ^cDivision of Surveillance, Hazard Evaluations and Field Studies, National Institute for Occupational Safety and Health (NIOSH), Cincinnati, OH, USA; ^dIllinois Fire Service Institute, University of Illinois at Urbana-Champaign, Champaign, IL, USA

ABSTRACT

A non-targeted analysis workflow was applied to analyze exhaled breath samples collected from firefighters pre- and post-structural fire suppression. Breath samples from firefighters functioning in attack and search positions were examined for target and non-target compounds in automated thermal desorption-GC/MS (ATD-GC/MS) selected ion monitoring (SIM)/scan mode and reviewed for prominent chemicals. Targeted chemicals included products of combustion such as benzene, toluene, xylenes, and polycyclic aromatic hydrocarbons (PAH) that serve as a standard assessment of exposure. Sixty unique chemical features representative of exogenous chemicals and endogenous compounds, including single-ring aromatics, polynuclear aromatic hydrocarbons, volatile sulfur-containing compounds, aldehydes, alkanes, and alkenes were identified using the non-targeted analysis workflow. Fifty-seven out of 60 non-targeted features changed by at least 50% from pre- to post-fire suppression activity in at least one subject, and 7 non-targeted features were found to exhibit significantly increased or decreased concentrations for all subjects as a group. This study is important for (1) alerting the firefighter community to potential new exposures, (2) expanding the current targeted list of toxicants, and (3) finding biomarkers of response to firefighting activity as reflected by changes in endogenous compounds. Data demonstrate that there are non-targeted compounds in firefighters' breath that are indicative of environmental exposure despite the use of protective gear, and this information may be further utilized to improve the effectiveness of personal protective equipment.



KEYWORDS

Non-targeted analysis; gas chromatography-mass spectrometry (GC/MS); automated thermal desorption (ATD); volatile organic compounds (VOC); breath research; firefighters


Introduction

Non-targeted analysis is a tool used for documenting environmental chemicals of exposure beyond those targeted for their adverse health effects. While many non-targeted methods were developed for liquid chromatography-mass spectrometry (LC/MS) applications (Newton et al. 2018; Rager et al. 2016), few techniques were developed for gas chromatography-mass spectrometry (GC/MS). As such, biomarkers in liquid media such as blood and urine have been more extensively studied than those in gas-phase exhaled breath. Wang et al. (2017) used GC/MS to identify non-targeted metabolites in human plasma, and Bergmann et al. (2018) developed a GC/MS method

for detecting semi-volatile compounds from environmental samples. Non-targeted analysis of breath-borne compounds may provide important information regarding environmental and occupational exposures to harmful compounds, including volatile organic compounds (VOC), semi-volatile organic compounds (SVOC), and polycyclic aromatic hydrocarbons (PAH). Non-targeted analysis is a rapidly evolving field requiring the harmonization of analytical instrumentation and data processing workflows to identify compounds from observed features (Milman and Zhurkovich 2017; Newton et al. 2018; Pleil and Stiegel 2013). In addition, compound libraries alone are often not sufficient for accurate compound identification. Information such as the number of data

CONTACT Joachim D. Pleil  pleil.joachim@epa.gov  Office of Research and Development, National Exposure Research Laboratory U.S. Environmental Protection Agency, Research Triangle Park, NC, USA

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uteh.

 Supplemental data for this article can be accessed [here](#).

© 2019 Taylor & Francis Group, LLC

sources, references in public databases, and retention time (RT) prediction may be utilized to narrow down the list of tentative compounds present in the samples (McEachran, Sobus, and Williams 2017; Sobus et al. 2017). In the current investigation, non-targeted analysis of firefighter breath samples was performed to identify non-targeted compounds that may indicate unknown exogenous exposures or changes in endogenous compounds due to firefighting activity. Although focused on occupational exposures, data demonstrate the general value of expanding public health exposure assessment to chemicals not previously found in environmental scenarios of interest.

Firefighters are afforded respiratory protection using self-contained breathing apparatus (SCBA) while turnout gear design is focused on dermal protection from thermal threats as well as cuts and abrasions, while providing some limited chemical exposure protection. Despite the use of personal protective equipment (PPE), several studies documented exposure and biological uptake of chemicals, such as benzene, benzene derivatives, and PAH during firefighting (Austin et al. 2001; Fent et al. 2014, 2015; Fernando et al. 2016; Laitinen et al. 2012; Navarro et al. 2017; Organtini et al. 2015; Pleil, Stiegel, and Fent 2014). Exposure of firefighters to VOC, PAH, and other harmful compounds has been assessed using breath sampling, personal air samples, skin wipes, urine, and blood (Fent et al. 2017; Fernando et al. 2016; Jayatilaka et al. 2017; Laitinen et al. 2014; Navarro et al. 2017; Stec et al. 2018). PAH, plasticizers, flame retardants, and particulate matter 2.5 μm ($\text{PM}_{2.5}$) were measured contaminating the surfaces of turnout gear, gloves, and hoods as well as infiltrating this protective layer to contaminate firefighters' skin (Alexander and Baxter 2016; Baxter et al. 2014; Easter, Lander, and Huston 2016; Fabian et al. 2014; Fent et al. 2017, 2015). VOC in breath and metabolites of PAH in urine were reported to increase in the min and hr after firefighting (Caux, O'Brien, and Viau 2002; Keir et al. 2017). These previous studies of building fires and firefighter exposures identified harmful compounds in air, dust, and biological media that have been subsequently utilized for targeted analysis. However, additional compounds identified through such studies may also become useful probes as biomarkers of chemical exposure and response beyond the firefighting environment.

This study describes the procedures and results for analyzing breath samples from a study of firefighters' exposures and physiological responses conducted by the University of Illinois Fire Service Institute (IFSI) with National Institute for Occupational Safety and Health (NIOSH) in collaboration with Underwriters Laboratories and the U.S. Environmental Protection Agency (EPA). Over 400 breath samples were collected and analyzed from firefighters responding to controlled residential fires. Fires were ignited in a 111 m^2 structure using common residential furnishings as the fuel for combustion, producing a variety of combustion products, including PAH, VOC, acid gases, hydrogen cyanide, and others (Fent et al. 2018; Horn et al. 2018). The primary targets for breath monitoring in this study were VOC and PAH. In the current investigation, non-targeted analyses of a subset of the firefighter breath samples ($n = 144$) were performed to identify additional compounds present in the samples that may be indicative of exposure. Biomarkers need not necessarily be the most abundant chemical present in the samples but should ideally show measurable differences between pre- and post-exposure states. Novel biomarkers may also be associated with acute or chronic health effects, supporting the need to analyze a broader spectrum of chemicals.

The analytical methods for analyzing the samples in this study were developed using automated thermal desorption-gas chromatography/mass spectrometry (ATD-GC/MS) with synchronous selected ion monitoring (SIM)/scan capabilities to target VOC, SVOC and PAH in SIM mode while also collecting the full scan total ion current (TIC) of the samples for non-targeted analysis of additional compounds (Geer Wallace et al. 2017). Using synchronous SIM/scan mode, few performance differences were reported in comparison to SIM- or scan-only methods for compound quantitation (Meng 2005; Strategies for Developing Optimal Synchronous SIM-Scan Acquisition Methods-AutoSIM/Scan Setup and Rapid SIM 2006). Further, a scan spectra may be employed for library searching to confirm identities of compounds targeted in SIM, which is required for some EPA methods (Strategies for Developing Optimal Synchronous SIM-Scan Acquisition Methods-AutoSIM/Scan Setup and

Rapid SIM 2006). While several SIM/scan methods were developed to improve and expand analyses of targeted compounds (Fahrenholtz et al. 2010; Van Thuyne, Van Eenoo, and Delbeke 2008; Wang et al. 2008), few demonstrations of SIM/scan have been shown for simultaneous analysis of targeted and non-targeted compounds. Gómez et al. (2009) used a GC/MS SIM/scan method to identify targeted compounds from personal care products as well as non-targeted organic contaminants in wastewater and river water. According to our knowledge, no apparent previous demonstrations of SIM/scan were identified using breath or other biological samples. Previously Geer Wallace et al. (2017) developed a SIM/scan method that exhibited reliable agreement between SIM and scan data for VOC calibration standards.

Here a comprehensive method is presented for non-targeted analysis of data collected using ATD-GC/MS and SIM/scan mode. Data and results for the investigation of breath samples collected from firefighters during controlled structure burns using the developed method are reported. Targeted chemicals were first identified and quantified in the SIM spectra, non-targeted features were discovered in the scan spectra using mass spectral fragments and RT, the NIST library was employed to tentatively identify the unknown features, and RT prediction software tools were tested to determine if the RT of the tentatively identified compounds matched the predicted RT of the compounds of interest. Non-targeted features displaying significant concentration differences between pre- and post-exposure concentrations were investigated as potential biomarkers.

Materials and methods

Study design

The study was conducted by the University of IFSI in collaboration with NIOSH. Study participants were recruited from across the nation and within IFSI, and firefighters who used tobacco, suffered from cardiovascular diseases, were older than 55 or pregnant were not permitted to participate in the study. Forty-one firefighters (37 male and 4 female) were selected. Four groups of 12 firefighters each participated in 4 controlled structure burns in residential settings using

identical structure layout and furnishings for fuel. Thirty-one of the firefighters participated in 4 controlled structure burns, 9 firefighters participated in two controlled structure burns, and one firefighter withdrew from the study due to illness. The groups participated in two burns on consecutive days the first week and another two burns on consecutive days the following week. Firefighting job assignments simulated typical fire-ground operations and included command/pump, attack (i.e., fire suppression), search and rescue, outside ventilation, and overhaul (backup and rapid intervention team (RIT)). Firefighters who went inside the structures (e.g., attack and search) were required to wear SCBA prior to entering the building, while firefighters participating outside the structure were allowed to select whether or not to wear SCBA (e.g., command/pump and outside ventilation) but were required to wear SCBA while working within the structure. More information about the study design and execution may be found in the literature (Fent et al. 2017; Horn et al. 2018).

Sample collection

Breath samples were collected at IFSI using a procedure approved by the University of Illinois and NIOSH Institutional Review Boards (IRB) with informed consent of all participants. NIOSH collected the exhaled breath samples utilizing a previously published sampling technique (Fent et al. 2013). Briefly, participants exhaled into BIO-VOC samplers, and 129 ml end-tidal breath was loaded onto Carbograph 2TD/1TD dual bed sorbent tubes (catalog no. C2-AXXX-5126, Markes International, Gold River, CA, USA). Breath samples were collected from firefighters before, immediately after, and 1 hr after participation in controlled structure burns (i.e., pre-, post-, and 1 hr post-exposure samples). Twenty field blanks were also collected at the fire-ground site, and concentrations of VOC in the room air inside the structure were measured (Fent et al. 2018). Participants were instructed to avoid secondhand smoke and chargrilled foods for 24-hr prior to and during the study. Urine samples that were collected from firefighters as part of this study (results not included) confirmed that cotinine levels were consistent with those of individuals that do not use tobacco. Meals were provided to firefighters to ensure that dietary restrictions were followed. Samples were

shipped to U.S. EPA on ice packs and stored at 4°C until ATD-GC/MS analysis.

ATD-GC/MS analysis

Details regarding standards preparation, calibration, and targeted analysis using SIM/scan mode may be found in Geer Wallace et al. (2017). Briefly, a 650 TurboMatrix ATD system (PerkinElmer LAS, Shelton, CT, USA) and a 6890N GC coupled to a 5975 inert XL MS (Agilent Technologies, Santa Clara, CA, USA) were utilized to desorb and analyze the samples. An Rxi-5Sil MS capillary GC column with 5 m integra guard, 30 m, 0.25 mm ID, 0.25 μ m film (part no. 13623-124, Restek Corporation, Bellefonte, PA, USA) was utilized for chromatographic separation. Thermal desorption was performed with a purge time of 5 min, the desorption flow rate of 20 ml/min, desorption time of 15 min, tube temperature of 375°C, and valve temperature of 270°C. The trap temperature was increased linearly from 10°C to 385°C with a trap hold of 10 min. The column flow and outlet splits were set to 2 ml/min and 6 ml/min, respectively. The oven program was as follows: 35°C for 2 min, 6°C/min to 190°C, and 28°C/min to 310°C with an 8-min hold. The quadrupole, ion source, and transfer line temperatures were 176, 290, and 290°C, respectively. Scan spectra were collected at a rate of 2² and the SIM/scan mode option was utilized, which allows for the collection of both target ion (SIM) and non-targeted (scan) spectra simultaneously (Strategies for Developing Optimal Synchronous SIM-Scan Acquisition Methods-AutoSIM/Scan Setup and Rapid SIM 2006). Research grade helium gas (99.9999%) and ultra-zero air were supplied from Airgas (Morrisville, NC, USA). The gaseous TO-14A 43 Component Mix at 1 ppm in nitrogen was purchased from Linde Electronics & Specialty Gases (Stewartville, NJ, USA).

Data analysis procedures

Chromatographic peak integrations were performed using ChemStation software version D.02.00. A five-tiered identification procedure was implemented to develop a master list of non-targeted compounds observed in the breath samples (Figure 1).

Scan chromatograms from 19 pre-, post-, and 1-hr post-exposure samples, 4 field blanks, and 2 method blanks were searched for non-targeted compounds, and the target and qualifier ions of each chromatographic peak were recorded to create a master list of features for investigation of non-targeted compounds. Sixty features present in 8 or more of the samples were selected for further investigation. Features that corresponded to column bleed (m/z 209 and 281) were excluded from the analysis. The pre-, post-, and 1-hr post-exposure samples for the 24 firefighters who participated in the attack and the 24 firefighters who participated in search and rescue positions (144 samples total) were investigated for the presence of these non-targeted features. The mass spectra of the chromatographic peaks were searched using the NIST library, and reverse match scores of 800 or higher were considered for potential matches. The mass spectra of unknown features were compared to the spectra of compounds matched using the NIST library searches. Subject matter expertise was used to narrow down the NIST library matches based upon the probability of detecting each compound in the sample at the observed RT based on molecular weight, stereochemistry, and elution order. Features with NIST identifications that met these criteria were recorded as tentative compound matches. Extracted ion chromatograms (EICs) of the top two most abundant ions were obtained, and peaks that appeared to contain more than one compound were assessed to identify compounds that co-eluted by separating out the target and qualifier ions of each feature.

Estimation of non-targeted compound concentrations

Using the target ions listed in Table 1, EICs of the non-targeted features were obtained for the pre-, post-, and 1-hr post-exposure samples collected from attack and search firefighters using Agilent MassHunter Qualitative Analysis Navigator software version B.08.00. Peaks were integrated to obtain the area under the curve, and the area counts were normalized using the toluene response for each sample. The toluene response was calculated as the ng of toluene in the sample divided by the area counts of the EIC for 91 m/z at 4.28 min in the SIM spectrum.

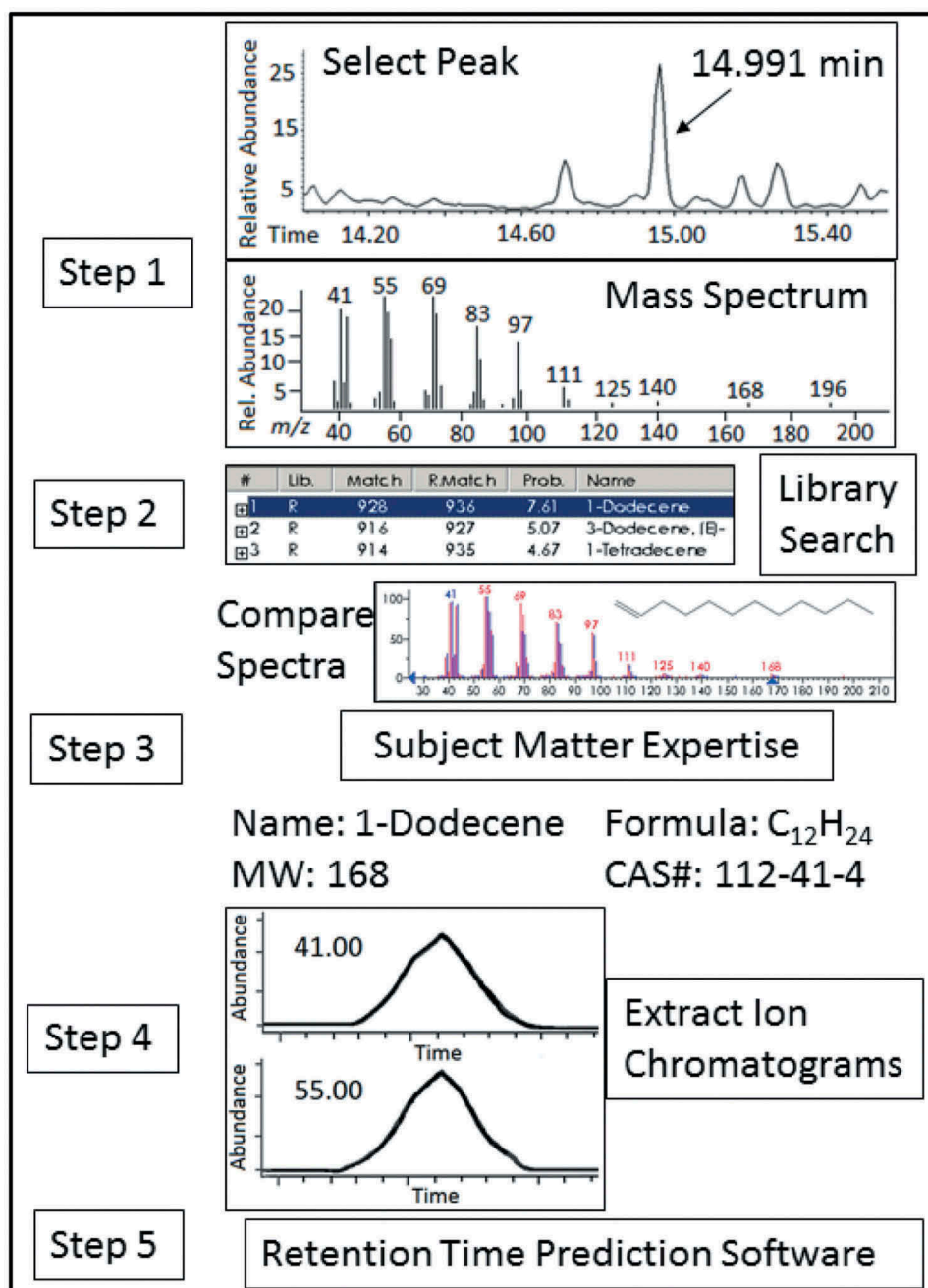


Figure 1. Non-targeted compound identification workflow. Peaks were selected using ChemStation software and mass spectra were compared to tentative matches in the NIST library. Subject matter expertise was utilized to determine the most likely identification, and the extracted ion chromatograms were obtained for the top two most abundant ions at that RT. For potential matches, RT prediction software was utilized as a verification step.

The toluene responses were multiplied by the area counts of the non-targeted compounds to obtain the apparent ng on tube with respect to toluene. Box-and-whisker plots of the pre-, post-, and 1-hr post-exposure concentrations for the attack and search firefighters were created in GraphPad Prism version 7 and are shown in \log_{10} scale from 5% to 95% without outliers for data visualization. Non-detects

were not included in graphing or statistical analyses. Unpaired two-tailed t-tests were performed in GraphPad Prism version 7 using non-transformed data to detect significant differences between pre- and post-exposure concentrations ($p < 0.05$).

The % change of the compound concentrations (ng/tube) from pre- to post-exposure was determined for all 48 attacks and search

Table 1. Non-targeted compounds identified in firefighter breath samples. Features without identifications listed were found in >8 samples, but a reasonable identity could not be obtained using the identification workflow. The compound names listed in this table are tentative identifications, and further validation is required to confirm their presence in the breath. Features in bold had significant changes in concentration from pre- to post-exposure (unpaired t-test, $p < 0.05$). VSC: Volatile sulfur-containing; PAH: polyaromatic hydrocarbon; PNA: polynuclear aromatic hydrocarbon; SRA: single ring aromatic; ID: identification; RT: retention time. No stars (weak match): ChromGenius RT error >10%; One star (moderate match): ChromGenius RT error 5–10%; Two stars (strong match): ChromGenius RT error <5% OR Restek RT error <5%; Three stars (very strong match): ChromGenius RT error <5% AND Restek RT <5% OR confirmed with standard (see Table S2 for ChromGenius and Restek RT parameters).

Feature No.	DTXSID	Preferred Name	Compound Class	Target Ion (m/z)	Qualifier Ion (m/z)	Formula	RT (min)
1	DTXSID8021482	Acetone**	Ketone	43	58	C ₃ H ₆ O	1.745
2	DTXSID3020833	Methyl tert-butyl ether**	Alcohol	59	41	C ₄ H ₁₀ O	1.829
3	DTXSID6023947	Carbon disulfide**	VSC	76	44	CS ₂	1.894
4	DTXSID8020204	tert-Butyl alcohol**	Ether	73	57	C ₅ H ₁₂ O	2.002
5	DTXSID5024394	Acetic acid**	Acid	43	45	CH ₃ COOH	2.096
6	DTXSID90190686	1,3,5-Trifluorobenzene**	SRA	132	81	C₆H₃F₃	2.353
	DTXSID30190140	1,2,4-Trifluorobenzene**					
7	DTXSID8026145	Thiophene**	VSC	84	58	C ₄ H ₄ S	2.768
8		No Match		86	147		2.892
9		No Match		41	56		3.169
10	DTXSID4020533	1,4-Dioxane**	SRA	88	58	C ₄ H ₈ O ₂	3.238
11	DTXSID4025117	Methyl disulfide**	VSC	94	79	C ₂ H ₆ S ₂	3.820
12		No Match		97	98		4.350
13	DTXSID8060666	Thiophene, 3-methyl-**	VSC	97	98	C ₅ H ₆ S	4.508
14	DTXSID7051769	2-Ethyl-1-hexene**	Alkene	70	55	C ₈ H ₁₆	4.750
15		No Match		229	73		5.081
16		No Match		189	194		5.288
17	DTXSID70883408	Cyclohexene, 3,3,5-trimethyl-**	SRA	109	67	C ₉ H ₁₆	5.599
	DTXSID00883618	Cyclohexene, 3,5,5-trimethyl-**					
18	DTXSID00276647	2-Propyl-1,3-dioxolane**	SRA	73	45	C ₆ H ₁₂ O ₂	6.212
19	DTXSID40214381	Phenol, 2,3-difluoro-**	SRA	130	82	F ₂ C ₆ H ₃ OH	7.000
	DTXSID50190142	Phenol, 2,4-difluoro-**					
20	DTXSID2059562	1-Nonene**	Alkene	43	56	C ₉ H ₁₈	7.266
21		No Match		209	181		8.580
22	DTXSID3042219	Propylbenzene**	SRA	91	120	C ₉ H ₁₂	8.775
23	DTXSID3035214	1-Octen-3-ol**	Alcohol	57	72	C ₈ H ₁₆ O	8.859
	DTXSID30865023	1-Nonen-3-ol**					
24	DTXSID8039241	Benzaldehyde**	SRA	105	106	C ₇ H ₆ O	8.972
25	DTXSID5021124	Phenol**	SRA	94	66	C ₆ H ₆ O	9.476
26	DTXSID7021491	Benzonitrile*	SRA	103	76	C ₇ H ₅ N	9.585
27	DTXSID6021402	1,2,4-Trimethylbenzene***	SRA	105	120	C₉H₁₂	9.890
28	DTXSID2060599	o-Vinyltoluene**	SRA	118	89	C ₉ H ₁₀	9.896
29	DTXSID6024913	Decane**	Alkane	57	43	C₁₀H₂₂	10.099
30	DTXSID8047769	1,2,3-Trimethylbenzene**	SRA	105	120	C ₉ H ₁₂	10.500
31		No Match		119	134		10.667
32		No Match		57	41		10.810
33	DTXSID90181594	Phenol, 2,5-difluoro-**	SRA	130	101	F ₂ C ₆ H ₄ O	11.082
	DTXSID00182456	Phenol, 2,6-difluoro-**					
	DTXSID10181596	Phenol, 3,5-difluoro-**					
34	DTXSID1021792	Salicylaldehyde**	SRA	122	121	C ₇ H ₆ O ₂	11.146
35	DTXSID6022472	Butylbenzene**	SRA	91	92	C ₁₀ H ₁₄	11.492
36	DTXSID6021828	Acetophenone***	SRA	105	77	C ₈ H ₈ O	11.735
37		No Match		180	161		12.140
38	DTXSID9021689	Undecane*	Alkane	57	43	C ₁₁ H ₂₄	12.698
39		No Match		267	73		13.721
40		No Match		200	169		13.785
41	DTXSID6022054	Pentylbenzene*	SRA	91	92	C ₁₁ H ₁₆	14.077
42	DTXSID5026914	1-Dodecene*	Alkene	43	55	C ₁₂ H ₂₄	14.991
43	DTXSID6027266	Tridecane*		57	43		15.200
44	DTXSID4021553	Decanal*	Aldehyde	55	70	C ₁₀ H ₂₀ O	15.299
45		No Match		112	140		15.558
46	DTXSID8061476	Benzene, hexyl-*	SRA	91	92	C ₁₂ H ₁₈	16.562
47		No Match		222	250		17.078
48		No Match		55	43		17.347
49	DTXSID4020878	2-Methylnaphthalene***	PAH	142	141	C ₁₁ H ₁₀	17.386
50	DTXSID1027267	Tetradecane		57	43		17.555

(Continued)

Table 1. (Continued).

Feature No.	DTXSID	Preferred Name	Compound Class	Target Ion (m/z)	Qualifier Ion (m/z)	Formula	RT (min)
51	DTXSID4020161	Biphenyl***	PNA	154	153	C ₁₂ H ₁₀	19.265
52	DTXSID1065421	1-Pentadecene	Alkene	69	55	C ₁₅ H ₃₀	19.583
53	DTXSID10881160	2-Methyl-1,1'-biphenyl*	PNA	168	167	C ₁₃ H ₁₂	19.632
54		No Match		57	43		19.762
55	DTXSID6021589	Dodecanal	Aldehyde	57	41	C ₁₂ H ₂₄ O	19.921
56		No Match		57	43		21.848
57	DTXSID00893333	Hexasulfur*	VSC	192	128	S ₆	21.858
58	DTXSID9047170	Nonadecane	Alkane	57	43	C ₁₉ H ₄₀	25.714
59	DTXSID0066377	1-Nonadecene	Alkene	69	41	C ₁₉ H ₃₈	27.374
60	DTXSID80872924	Octasulfur**	VSC	64	256	S ₈	30.114

firefighters. Pre- and post-exposure samples were paired for the firefighters, and only compounds that were identified and quantified in both samples were included in the analysis. Estimated compound concentrations that were less 0.1 ng/tube were excluded (lowest 7.5% of the data). The differences between the pre- to post-exposure concentrations of each non-targeted compound were calculated, and % change in the compound concentrations was determined by dividing this difference by the pre-exposure concentration and multiplying by 100. Percent changes that were greater than 50% (or <-50%) were considered significant increases or decreases in post-exposure concentrations, respectively, at the individual level.

Retention time (RT) prediction

Retention time prediction was employed as the last step in Figure 1. The Restek EZGC chromatogram modeler was used for method development to estimate RT of compounds present in the Restek library based upon the analytical column and GC/MS parameters utilized in the analysis. NIST library compound identifications with RT that matched those predicted using the same Restek column were considered more confident matches.

Chromatography data were modeled using Advanced Chemistry Development, Inc (ACD)/Labs software ACD/ChromGenius (Toronto, Ontario, Canada) to predict RT for candidate compounds assigned as features. The model was developed by dividing compounds from the standards with known RT into a training set of 34 chemicals (~80%) and a test set of 8 chemicals (~20%) (Table S1). These chemicals originated from either the Linde TO-14 A 43 component

mix or the Restek PAH standard and were previously analyzed using the same column and chromatographic conditions as employed in the current method. In the similarity search options tab, the dice coefficient was selected to show the best 25 records. The equation was correlated linearly with 4 points per parameter and 0.01 delta error for the new parameter in the equation. The equation was set to not be utilized if the error was >0.9. The following GC parameters were used for modeling: boiling point, octanol-water partitioning coefficient of the molecule (logP), polar surface area, molecule volume, molecule weight, molar refractivity, H donors, and H acceptors.

Before applying the RT prediction model to the unknown compounds, the model was evaluated using 4 different criteria. As established previously as a useful criterion for assessing the predictive ability of RT models (Aalizadeh et al. 2016; McEachran et al. 2018), the coefficient of determination (R^2) between the predicted and experimental RT was calculated using Eqn. 1:

$$R^2 = 1 - \frac{\sum_{i=1}^{n_{TR}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_{TR}} (y_i - \bar{y})^2} \quad (1)$$

where y_i and \hat{y}_i are the experimental and estimated responses, respectively, \bar{y} is the mean response, and n_{TR} is the number of compounds in the training set.

The root mean squared error (RMSE) between the experimental and estimated RT for the training compounds was calculated using Eqn. 2:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2)$$

where y_i and \hat{y}_i are the experimental and estimated RTs, respectively. The model was also evaluated

based on the absolute errors (differences) of the experimental versus predicted RT and the percentages of the chromatographic run time (40.12 min) that these values fell within, with smaller percentages indicating more accurate matches.

The developed model was applied to 49 compounds that were identified by the NIST library as being tentatively present in the firefighter samples. Some of these compounds were isomers that corresponded to the same feature. To create the model, the GC column, oven program, and other method parameters were input into the software and used to predict RT of identified compounds. Additional parameters and equations may be found in the Supporting Information.

Results and discussion

Identification of endogenous and exogenous non-targeted compounds in breath samples

The procedure outlined in Figure 1 was utilized to generate a list of 60 prevalent non-targeted features present in the pre-, post-, and 1-hr post-exposure breath samples as well as method and field blanks, as presented in Table 1. Features without reasonable tentative identifications are left blank. Additional information regarding the identified feature, including CAS-RN, SMILES, as well as ChromGenius and Restek EZGC Chromatogram modeler predicted RT are found in Table S1.

Figure 2 depicts the abundance of each identified compound class among the 60 features listed in Table 1. The breath samples appear to be largely composed of single-ring aromatics (SRA), volatile sulfur-containing compounds (VSC) and alkenes. Some of these compounds appear to be exogenous contaminants from chemical exposure, while others may be metabolites that resulted from an endogenous response to exposure. While it might be difficult to distinguish between some endogenous and exogenous VOC (Sun, Shao, and Wang 2016), the presence of certain compounds in the post- versus pre-exposure samples might be used to discern the origins of some of the observed non-targeted compounds.

Compounds detected in the pre-exposure breath samples that were not present in the method or field blanks were presumed to originate from the

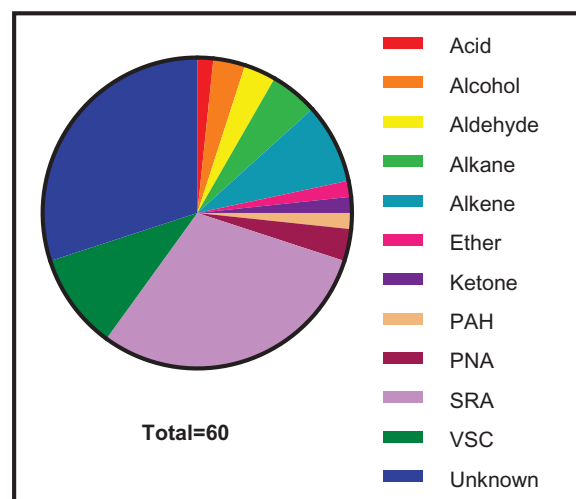


Figure 2. Classification of non-targeted compounds identified in firefighter samples ($n = 40$). Compounds were divided into classes based on compound functional groups to determine the distribution of compound classes within the breath samples. PAH: polycyclic aromatic hydrocarbon; PNA: polynuclear aromatic hydrocarbon; SRA: single ring aromatic; VSC: Volatile sulfur-containing.

firefighters' breath, and might consist predominantly of endogenous compounds. Endogenous compounds in breath are produced through biological processes and metabolic pathways within the body, some of which may result from oxidative stress (Amann et al. 2014; Phillips et al. 1999). Endogenous compounds such as acetone and alcohols are a signature of cellular and metabolic processes and were found to be present in almost all breath samples, regardless of exposure (Filipiak et al. 2016; Hakim et al. 2012; O'hara et al. 2009). Other compounds detected in pre-exposure breath samples included carbon disulfide, 1,4-dioxane, thiophene, dimethyl disulfide, 3-methyl thiophene, biphenyl, benzonitrile, and acetophenone may represent exposures that the firefighters encountered prior to beginning the exercise. Straight chain and branched alkanes as well as benzene and naphthalene derivatives were also detected, and these compounds were previously observed in the breath of unexposed individuals (Phillips et al. 1999). Low levels of some of these compounds were detected in the pre-exposure breath samples, indicating that these chemicals may be of endogenous origins. While firefighters may be predisposed to a higher baseline of VOC than the average population, additional exposures prior to sampling, such as filling up a tank with gasoline or walking past a person smoking a cigarette, might

impact pre-exposure breath profiles (Hakim et al. 2012; Jo and Song 2001; Lindstrom and Pleil 1996). Off-gassing of chemical contaminants from firefighter PPE was also identified as an exposure route (Fent et al. 2015; Kirk and Logan 2015). Because firefighters in this study were around their PPE and other equipment before firefighting, some exposure may have occurred prior to the pre-exposure breath sampling. Although most of the VOC contaminants from prior use (previous day) would have dissipated by this time, some exposure to SVOC may have occurred (Fent et al. 2018).

Some VOC identified in the samples appear to be indicative of an endogenous response to chemical exposure from firefighting. The endogenous response may be influenced by the heavy metabolic demands that the firefighters experienced during firefighting exercises. For example, aldehyde production was previously shown to occur as a response to smoke and indoor air pollution (Filipiak et al. 2016). The production of reactive oxygen species (ROS) may be triggered by exposure to cigarette smoke and pollution (Hakim et al. 2012). ROS were found to lead to increased hydrocarbon production, which was noted in the post-exposure samples. Peroxidation of polyunsaturated fatty acids may also result in an elevation in straight chain alkanes, such as heptane, octane, and decane (Filipiak et al. 2016; Hakim et al. 2012). Oxidation of alkanes, influenced by the presence of aromatic hydrocarbons from tobacco smoke, was also shown to increase alcohol, ketone, and aldehyde production (Filipiak et al. 2016). These biological processes may explain how non-targeted compounds such as decane, undecane, tridecane, decanal, tetradecane, dodecanal, and nonadecane may have been generated in response to the firefighting activities.

Exogenous compounds are a direct indicator of exposure and often enter the body through inhalation, dermal absorption, or ingestion (Bos, Sterk, and Fowler 2016; Calenic and Amann 2014). Aromatic compounds such as benzene, toluene, and styrene as well as nitriles are considered to be exogenous compounds that predominantly originate from smoke, environmental contaminants, gasoline, and vehicle emissions (Filipiak et al. 2016). Consistent occupational exposure to products of combustion may enable these aromatic compounds to collect within fatty tissues, leading to the slow release of these compounds in the breath over time (Hakim et al. 2012).

As such, repeated exposure of firefighters to aromatic compounds prior to this study may explain the presence of these compounds in high levels even in the pre-exposure breath samples. 2-Methylnaphthalene is a PAH that is a byproduct of combustion and emitted by petroleum products, jet fuel, and coal tar creosote (Kim, Andersen, and Nylander-French 2006; Lin et al. 2009). While the toxicity of methylnaphthalenes was investigated in animal models, the long-term effects of inhalation exposure to these compounds are not well documented in humans (Lin et al. 2009). As illustrated in Table 1, several fluorinated compounds were also detected, including trifluorobenzenes and several difluorophenols. Fluorinated compounds are not endogenous and might indicate exogenous exposure to environmental contaminants or possible contact with fluorinated fire suppression agents (outside the scope of this study) (Shimoda and Hamada 2010). The presence of benzene derivatives in post-exposure samples might be both due to exogenous exposure and endogenous response. These compounds may have been directly generated as byproducts of combustion or may be due to endogenous metabolism of benzene exposure at the fire-ground. The differences in the number and structures of these derivatives observed between individuals show that the metabolic processing of these VOC may vary significantly from individual to individual.

Retention time prediction using ACD/ChromGenius and Restek EZGC chromatogram modeler

Evaluation of the ACD/ChromGenius model

As an example of the utility of RT prediction models, ACD/ChromGenius was used to develop a model to predict the RT of tentative compound identifications that were assigned to non-targeted features using the NIST library. This software program uses information based upon the quantitative structure–property relationship (QSPR) and quantitative structure–retention relationship (QSRR) modeling to predict RT of compounds (Aalizadeh et al. 2016; Bade et al. 2015; Dossin et al. 2016; Veenaas, Linusson, and Haglund 2018). Properties of the compounds such as octanol-water partition coefficient (logP), pH-dependent partition coefficient (logD), molecular mass, and polarity are utilized by the software to aid in RT prediction (Dossin et al. 2016; McEachran et al. 2018). The

model was developed using RT provided for a known set of standards or chemicals to predict the elution characteristics of non-targeted compounds. ACD/ChromGenius was previously utilized to predict RT of VOC from tobacco aerosol analyzed by GC-high resolution-MS with accuracy between 85% and 115% (Dossin et al. 2016), and ACD/ChromGenius was employed to build RT prediction models using GC×GC-MS data for a variety of chemical classes (Veenaas, Linusson, and Haglund 2018).

The model was employed to determine if the predicted RT for the non-targeted compound IDs matched the RT of features detected in the samples to gain confidence in the identifications. The model was developed utilizing training and test set chemicals that originated from the standards analyzed by ATD-GC/MS (Table S2). A limitation of the ChromGenius method development is that only standards analyzed using the same column and chromatographic conditions as the firefighter samples might be employed to compose training and test sets. Since the SIM/scan method was developed to target benzene derivatives and PAH, most of the compounds in the standards contained carbon, hydrogen, chlorine, fluorine, and bromine, but did not include oxygen. This limited the applicability of the model to some of the non-targeted compounds that were observed, such as the aldehydes, ketones, and alcohols, although these compound classes were not largely represented in this study. Most of the non-targeted compounds tested for RT prediction were single-ring aromatics, which were well represented in the training and test sets.

The model was evaluated using three different statistical parameters: the coefficient of determination between the predicted and experimental RT (R^2), RMSE, and average absolute mean error. The R^2 value equaled 0.75 for the training set (set 1), test set (set 2), and combined set (set 3) (Table 2). The R^2 obtained for this model exhibited reasonable agreement between predicted and experimental RT considering the small sizes of the training and test sets and is within the range of R^2 values previously obtained for similar retention time prediction models (McEachran et al. 2018). The RMSE was highest for the test set (set 2), which only contained 8 standards, while the training and combined sets (sets 1 and 3) were larger and thus

had lower RMSE. The absolute mean error (AME), or the average difference between the predicted and experimental RT, was low for all three sets.

Each compound in sets 1–3 was also assessed for RT error using RT windows, represented as a fractional percentage of the chromatographic run time, as a metric to assess how close the predicted RT matched to experimental RT. Previously McEachran et al. (2018) developed ACD/ChromGenius method to predict the RT of 97% of compounds in the training set to be within a RT window that spanned 20% of the chromatographic run time, and only 46% of the compounds in the training set had predicted RT within a 5% window of the chromatographic run time. While this study had fewer compounds in the training set, a higher % of compounds in the training and test sets had predicted RT that fell within a 5% window of the experimental RT (2.006 min). In set 1, 97% of the compounds in the training set had RT within the 5% window compared to 75% of compounds in set 2 and 93% of compounds in set 3. The only three compounds found to be outside of the 5% RT window were phenanthrene in set 1 (2.16 min), fluorene and anthracene in set 2 (2.33 and 2.35 min, respectively), and all three of these compounds in set 3. These compounds are PAH with later RT (between 23.47 and 27.31 min). This may indicate that the model is slightly more accurate at predicting RT of lighter compounds with earlier RT, which is reasonable considering that most of the compounds in the test set had RT of less than 15 min. Predictive accuracy might potentially be improved by increasing the size of the training set and including more compounds with later RT, but in comparison to previously published methods, the error between the predicted and experimental RT in this method is low (McEachran et al. 2018)

Table 2. Summary statistics for the three compound sets evaluated using ACD/ChromGenius RT prediction.

Set Name	<i>n</i>	R^2	RMSE (min)	AME (min)	5% ^a
1) Training	34	0.75	0.43	0.77	33
2) Test	8	0.75	0.91	1.08	6
3) Training + Test	42	0.75	0.52	0.83	39

n number of chemicals in the set, R^2 coefficient of determination, RMSE root mean squared error, AME absolute mean error.

^aThe number of compounds with predicted RTs falling within a 5% window of the total chromatographic run time.

Application of the ACD/ChromGenius and Restek models for RT prediction

The ChromGenius model was applied to predict RT of 49 non-targeted compounds identified by the NIST library that corresponded to 44 detected features. Some features matched well to multiple compounds in the NIST library due to the potential for isomers or compounds with similar formulas. The predicted RT of these compounds were compared to experimental RT of the corresponding features, and the AMEs and run time % were calculated to evaluate the strength of the NIST library matches based upon the predicted chromatography. Results of the ChromGenius RT predictions are found in Table S2. The matches in Table S2 were ranked according to the performance of the training and test sets, in which 93–97% of compounds had predicted RT within a 5% window of the chromatographic run time.

Thirty-six of the unknown candidate compounds had predicted RT differences that fell within a window of less than 5% of the chromatographic run time for the detected features for either ChromGenius or EZGC predictions (strong matches), and an additional 9 compounds displayed RT differences between 5% and 10% of the run time (moderate matches). These 45 compounds with RT differences of less than 10% of the run time were considered tentative feature identifications. The remaining 5 compounds that were evaluated (1-pentadecene, dodecanal, 1-nonadecene, tetradecane, and nonadecane) contained errors greater than 10% and were considered unlikely identifications (weak matches) according to the ChromGenius method (see Table S2). All of these compounds with high RT errors eluted at 17 min or later, in the portion of the chromatogram where few standards were available for the creation of the training and test sets during method development. Therefore, it is unsurprising that these compounds showed the highest RT error.

Some compounds with higher RT errors based upon the ChromGenius model were available for comparison using the Restek EZGC chromatogram modeler. In EZGC, the column, oven program, and flow rate used in the method were input in the program, and RT for compounds in the Restek database were estimated. While not all compounds of interest were available for

comparisons, several compounds performed better in the EZGC model, including benzaldehyde, phenol, decane, acetophenone, 2-methylnaphthalene, and biphenyl, which all had absolute RT errors less than 0.15 min using the EZGC predictions. Most of these compounds had absolute RT errors between 1 and 3 min in the ChromGenius model. Further, compounds with RT error less than 5% in the ChromGenius model that were further confirmed with RT error less than 5% using the EZGC model were labeled as very strong matches (8% of all matches).

The implementation of ACD/ChromGenius and Restek EZGC RT prediction in this study enhanced confidence in 90% of the NIST library assignments of unknown features. For features with multiple tentative identifications, RT modeling helped with feature assignment. For example, feature 6 was tentatively identified as either 1,3,5-trifluorobenzene or 1,2,4-trifluorobenzene. However, the absolute RT error for 1,3,5-trifluorobenzene was only 0.25 min while the error was 1.43 min for 1,2,4-trifluorobenzene. Therefore, the better compound assignment for this feature is 1,3,5-trifluorobenzene. Similarly, feature 23 was identified as either 1-octen-3-ol or 1-nonen-3-ol, but the absolute RT error for 1-octen-3-ol was 0.15 min while 1-nonen-3-ol was 1.59 min, making 1-octen-3-ol a more likely compound assignment (Table S2). When deciding which non-targeted compounds to pursue as additional biomarkers in future studies, RT prediction tools might help to assign the most probable compound identifications. While the ChromGenius model exhibited some error for later eluting compounds in model development, obtaining approximate RT estimates for feature assignments might add another level of confidence to library identifications, as compounds with high RT error may be incorrect matches.

Changes in non-targeted compound concentrations in post-exposure breath samples

The pre-, post-, and 1-hr post-exposure samples were assessed for significant changes in concentrations of the 60 non-targeted compounds from attack and search firefighters. The relative concentrations of the compounds were estimated according to the toluene response. This calculation technique was not meant to be fully quantitative but was used to provide

concentration estimates. Due to the implementation of SIM/scan, approximately 4–5 data points were collected per peak in scan, compared to 8–12 data points in SIM. Thus, in SIM/scan mode the scan spectra are considered to be more qualitative instead of quantitative (Strategies for Developing Optimal Synchronous SIM-Scan Acquisition Methods-AutoSIM/Scan Setup and Rapid SIM 2006).

Unpaired two-tailed t-tests ($\alpha = 0.05$) were performed to determine if there were significant differences between the pre- and post-exposure concentrations of all non-targeted compounds. Seven non-targeted features demonstrated significant differences between pre- and post-exposure concentrations. The normalized data are illustrated as box-and-whisker plots in Figure 3. Two features were elevated post exposure: feature 29 (decane) and feature 54 (m/z 57), while 5 features showed decreased post-exposure concentrations: feature 6 (trifluorobenzene), feature 27 (1,2,4-trimethylbenzene), feature 39 (m/z 267), feature 45 (m/z 112), and feature 47 (m/z 222). The identifications of several of these features were unable to be verified during data analysis, but moderate matches were obtained using the NIST library. The highest scoring identifications for feature 54 were 4-dimethyldecane, 2,2,5-trimethyl-3,4-hexanedione, and 3,3-dimethylhexane; for feature 39, 2TMS derivative 2,4-dihydroxybenzaldehyde; for feature 45, 2,2-diethyl-N-ethylpiperidine and 1-ethoxy-4-fluorobenzene; for feature 47, no probable matches were found. Further study will be required to verify the identities and behavior of these non-targeted compounds. While features 27, 6, 45, and 47 (A, B, F and G in Figure 3) also displayed significant differences between pre- and 1-hr post-exposure concentrations (unpaired two-tailed t-tests ($\alpha = 0.05$)), the 1-hr post-exposure time point represents the recovery period, and this finding does not impact the results of this study.

Based on the data in Figure 3, decane and feature 54 appear to be potential indicators of exposure for future investigation. Increased decane levels may indicate an endogenous response to exposure, as elevated production of straight chain alkanes might be produced by smoke exposure (Hakim et al. 2012). Although no direct link was made to this study, in addition to benzene, decane levels in breath were previously reported as a biomarker of lung cancer

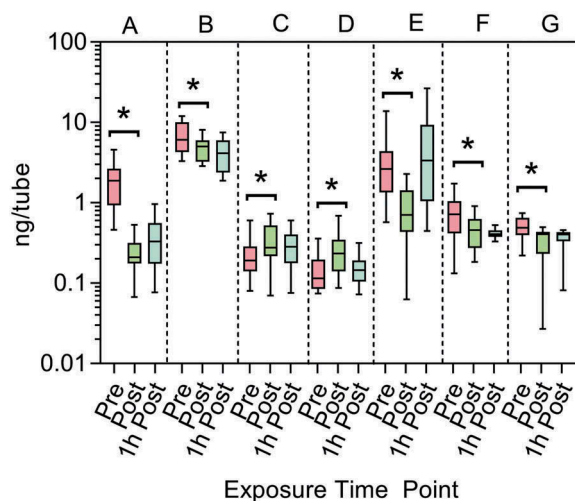


Figure 3. Concentrations (ng/tube) of selected non-targeted compounds measured in breath samples from attack and search firefighters. A) 1,2,4-Trimethylbenzene (feature 27, m/z 105); B) trifluorobenzene (feature 6, m/z 132); C) decane, (feature 29, m/z 57); D) feature 54 (m/z 57); E) feature 39 (m/z 267); F) feature 45 (m/z 112); G) feature 47 (m/z 222). Concentrations are shown as log₁₀ transformed intervals from 5% to 95% without outliers. Features were detected in the following number of pre-, post-, and 1-h post-exposure samples, respectively: A) $n = 37, 26, 31$; B) $n = 48, 47, 48$; C) $n = 27, 25, 33$; D) $n = 45, 43, 48$; E) $n = 23, 15, 9$; F) $n = 28, 21, 19$; G) $n = 19, 33, 29$. All compounds showed significant differences between pre- and post-exposure concentrations according to two-tailed t-tests ($p < 0.05$).

(Sun, Shao, and Wang 2016). While benzene was included as a targeted compound in this and other studies, decane is an interesting potential biomarker that might be targeted for exposure monitoring in future studies. The concentration of decane in the attack and search samples provides an indication of the potential levels of this compound in other firefighter samples, as firefighters who participated in the attack and search and rescue positions likely experienced the highest exposure in this study since these individuals spent the most time inside the building while contaminant concentrations were elevated. However, other fire-ground job assignments require firefighters to work near the building on the outside, where SCBA usage is not as consistent, which may also increase post-exposure breath concentration of non-targeted compounds. This is also important to note for wildland firefighters, who are exposed to high levels of air pollutants in wood smoke without the use of respiratory protection (Broyles 2013; Miranda et al. 2012; Navarro et al. 2017).

Although two non-targeted compounds were elevated post-exposure, the other 5 agents decreased in concentration, which may indicate that the firefighting gear is working well to protect against exposure to these chemicals or that individuals have variable metabolism and response to these compounds. The % changes of the feature concentrations from pre- to post-exposure were investigated for individual firefighters throughout the course of a single controlled burn exercise. The % changes are depicted in a heat map in Figure 4. As shown in Figure 4, all attack and search firefighters (x-axis) had a least one compound with a substantial increase ($>50\%$) or decrease ($<-50\%$) in post-exposure concentrations (approximately 44% of the percent changes). Fifty-seven out of the 60 non-targeted features also changed by at least $\pm 50\%$ from pre- to post-fire suppression activity in at least one subject, as shown in Figure 4.

Many firefighters showed trends in either increasing or decreasing post-exposure feature concentrations. Approximately one-third of the firefighters exhibited greater than 75% of their significantly changed concentrations as post-exposure increases, while a third showed primarily decreased post-exposure concentrations ($>75\%$), and the last third did not show a preference for elevated or reduced post-exposure concentrations (between 25% and 75%). The heat map also reveals the features that were found to be more common in the firefighters' breath. Two compounds, features 20 and 59, were not prominent amongst the paired firefighter samples, although these compounds were present in

unpaired breath samples, revealing that these features are probably not noteworthy.

The analysis of individual firefighters according to the heat map demonstrates that not all firefighters are experiencing significant exposure to the same compounds. This might be attributed to differences in firefighters' tasks and behavior during the controlled structure burns, which may result in unique exposure. For example, slight variations in the effectiveness and level of protection afforded by the gear may also lead to higher or lower exposure for individual firefighters. The heat map analysis may be utilized to postulate those individuals who may have less effective protective gear, which may subsequently be investigated to mitigate future exposures.

Conclusions and future directions

Non-targeted analysis was found to be an important complementary approach to standard targeted evaluations of compounds expected to exert adverse effects. This investigation demonstrates how to complement standard longitudinal exposure evaluations with a discovery component and concludes the following:

- Using SIM/scan mode during GC/MS analysis enables the documentation of both targeted and non-targeted compounds simultaneously, thus saving analytical effort.

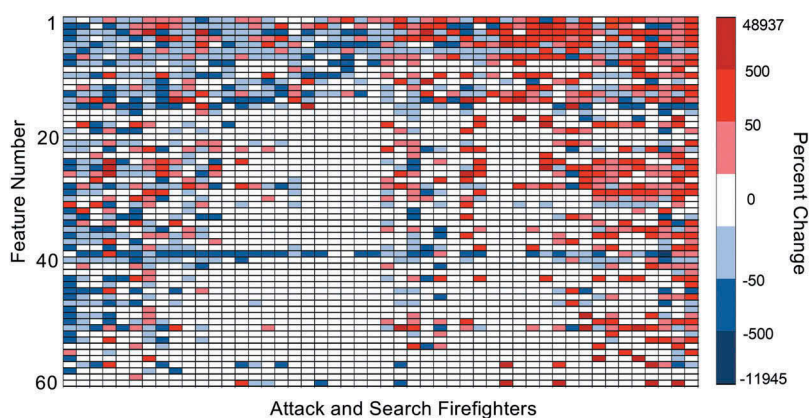


Figure 4. Heat map of percent changes of non-targeted features from pre- to post-exposure identified in attack and search firefighters. The 48 attack and search firefighters are shown on the x-axis and the features on the y-axis. Percent changes were considered substantial at $>50\%$ and $<-50\%$. Blank squares ($n = 1695$) show features that were not detected in both the pre- and post-exposure breath samples for that exposure event. The percent change ranges for each color category were counted as follows: 0 to 50 ($n = 252$), 50 to 500 ($n = 256$), 500 to 48937 ($n = 34$), ≥ 50 to 0 ($n = 401$), ≥ 500 to ≥ 50 ($n = 256$), ≥ 11945 to ≥ 500 ($n = 1$).

- Non-targeted chromatographic features might be interpreted as novel biomarkers for exposure and pre-clinical effect using longitudinal comparisons; it is not necessary to unambiguously identify the actual compounds as long as retention time and mass fragments are known.
- Grouped analyses, e.g., before vs. after firefighting activity, may not have statistical significance due to variability at the individual subject level; however, a heat-map approach can be used to identify those individuals more prone to exposures.

Data demonstrated how probative biomarkers that were not intentionally monitored might be identified for future assessments. Sixty prominent non-targeted features were detected in the pre- and post-exposure breath samples from firefighters participating in attack and search positions. Through comparative estimates, 7 non-targeted features were found to display significantly different concentrations from pre- to post-exposure (Table 1, Figure 3). Of these, decane was noted to have the highest significantly increased concentrations in post-exposure breath samples, making this compound an interesting probe for future studies. Other selected non-targeted compounds detected herein may also become targeted compounds in future firefighter studies to confirm compound assignments and to collect quantitative data for exposure assessment.

Ultimately, the information developed here may be used to refine future exposure assessments to identify specific activities, gear configurations, and behaviors that might affect environmental protection. Specific compounds, as deduced by the heat-map analysis, may be employed to (1) identify groups of individuals at greater risk, and (2) improve the effectiveness of protective gear. Future studies need to include targeting several of the identified non-targeted features to confirm identities and quantify levels of these compounds in firefighter breath during controlled structure burns. Additional types of sorbent materials may also be investigated to enhance the range of compounds that can be analyzed using this technique.

Acknowledgments

The authors thank the volunteer subjects who provided breath samples. Individuals were compensated up to \$599 to participate in this study. This study was approved by the NIOSH and University of Illinois Institutional Review Boards and was funded by the U.S. Department of Homeland Security Assistance to Firefighters Grant Fire Prevention & Safety program (EMW-2013-FP-00766). This study was also made possible through a partnership with the CDC Foundation. Dr. Sibel Mentese is grateful for the travel grant received from TUBITAK. The authors acknowledge Dr. Andrew McEachran and Dr. Jonathan Mosley from U.S. EPA for assisting with ACD/ChromGenius model development. This research has been subjected to EPA review and approved for publication. The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of NIOSH or EPA. Mention of trade names and commercial products does not constitute endorsement or recommendation for use.

Conflict of interest

The authors declare no competing financial interest.

ORCID

M. Ariel Geer Wallace  <http://orcid.org/0000-0002-3966-5101>

References

- Aalizadeh, R., N. S. Thomaidis, A. A. Bletsou, and P. Gago-Ferrero. 2016. Quantitative structure–Retention relationship models to support nontarget high-resolution mass spectrometric screening of emerging contaminants in environmental samples. *J. Chem. Inf. Model* 56:1384–98. doi:10.1021/acs.jcim.5b00752.
- Alexander, B. M., and C. S. Baxter. 2016. Flame-retardant contamination of firefighter personal protective clothing—A potential health risk for firefighters. *J. Occup. Environ. Hyg.* 13:D148–D155. doi:10.1080/15459624.2016.1183016.
- Amann, A., B. de Lacy Costello, W. Miekisch, J. Schubert, B. Buszewski, J. Pleil, N. Ratcliffe, and T. Risby. 2014. The human volatilome: Volatile organic compounds (VOCs) in exhaled breath, skin emanations, urine, feces and saliva. *J. Breath Res.* 8:034001. doi:10.1088/1752-7155/8/3/034001.
- Austin, C. C., D. Wang, D. J. Ecobichon, and G. Dussault. 2001. Characterization of volatile organic compounds in smoke at municipal structural fires. *J. Toxicol. Environ. Health Part A* 63:437–58. doi:10.1080/152873901300343470.
- Bade, R., L. Bijlsma, T. H. Miller, L. P. Barron, J. V. Sancho, and F. Hernández. 2015. Suspect screening of large

- numbers of emerging contaminants in environmental waters using artificial neural networks for chromatographic retention time prediction and high resolution mass spectrometry data analysis. *Sci. Total Environ.* 538:934–41. doi:10.1016/j.scitotenv.2015.08.078.
- Baxter, C. S., J. D. Hoffman, M. J. Knipp, T. Reponen, and E. N. Haynes. 2014. Exposure of firefighters to particulates and polycyclic aromatic hydrocarbons. *J. Occup. Environ. Hyg.* 11:D85–D91. doi:10.1080/15459624.2014.890286.
- Bergmann, A. J., G. L. Points, R. P. Scott, G. Wilson, and K. A. Anderson. 2018. Development of quantitative screen for 1550 chemicals with GC-MS. *Anal. Bioanal. Chem.* 410:3101–10. doi:10.1007/s00216-018-0997-7.
- Bos, L. D., P. J. Sterk, and S. J. Fowler. 2016. Breathomics in the setting of asthma and chronic obstructive pulmonary disease. *J. Allergy Clin. Immunol.* 138:970–76. doi:10.1016/j.jaci.2016.08.004.
- Broyles, G. A. 2013. Wildland firefighter smoke exposure study. *Utah State University: All Graduate Plan B and other Reports.* 356. <https://digitalcommons.usu.edu/gradreports/356>.
- Calenic, B., and A. Amann. 2014. Detection of volatile malodorous compounds in breath: Current analytical techniques and implications in human disease. *Bioanalysis* 6:357–76. doi:10.4155/bio.13.306.
- Caux, C., C. O'Brien, and C. Viau. 2002. Determination of firefighter exposure to polycyclic aromatic hydrocarbons and benzene during fire fighting using measurement of biological indicators. *Appl. Occup. Environ. Hyg.* 17:379–86. doi:10.1080/10473220252864987.
- Dossin, E., E. Martin, P. Diana, A. Castellon, A. Monge, P. Pospisil, M. Bentley, and P. A. Guy. 2016. Prediction models of retention indices for increased confidence in structural elucidation during complex matrix analysis: Application to gas chromatography coupled with high-resolution mass spectrometry. *Anal. Chem.* 88:7539–47. doi:10.1021/acs.analchem.6b00868.
- Easter, E., D. Lander, and T. Huston. 2016. Risk assessment of soils identified on firefighter turnout gear. *J. Occup. Environ. Hyg.* 13:647–57. doi:10.1080/15459624.2016.1165823.
- Fabian, T. Z., J. L. Borgerson, P. D. Gandhi, C. S. Baxter, C. S. Ross, J. E. Lockey, and J. M. Dalton. 2014. Characterization of firefighter smoke exposure. *Fire Technol.* 50:993–1019. doi:10.1007/s10694-011-0212-2.
- Fahrenholtz, S., H. Hühnerfuss, X. Baur, and L. T. Budnik. 2010. Determination of phosphine and other fumigants in air samples by thermal desorption and 2D heart-cutting gas chromatography with synchronous SIM/Scan mass spectrometry and flame photometric detection. *J. Chromatogr. A* 1217:8298–307. doi:10.1016/j.chroma.2010.10.085.
- Fent, K. W., B. Alexander, J. Roberts, S. Robertson, C. Toennis, D. Sammons, S. Bertke, S. Kerber, D. Smith, and G. Horn. 2017. Contamination of firefighter personal protective equipment and skin and the effectiveness of decontamination procedures. *J. Occup. Environ. Hyg.* 14:801–14. doi:10.1080/15459624.2017.1334904.
- Fent, K. W., J. Eisenberg, D. Evans, D. Sammons, S. Robertson, C. Striley, J. Snawder, C. Mueller, V. Kochenderfer, and J. Pleil. 2013. Evaluation of dermal exposure to polycyclic aromatic hydrocarbons in fire fighters. Health Hazard Evaluation Report (2010–0156–3196).
- Fent, K. W., J. Eisenberg, J. Snawder, D. Sammons, J. D. Pleil, M. A. Stiegel, C. Mueller, G. P. Horn, and J. Dalton. 2014. Systemic exposure to PAHs and benzene in firefighters suppressing controlled structure fires. *Ann. Occup. Hyg.* 58:830–45. doi:10.1093/annhyg/meu036.
- Fent, K. W., D. Evans, K. Babik, C. Striley, S. Bertke, S. Kerber, D. Smith, and G. Horn. 2018. Airborne contaminants during controlled residential fires. *J. Occup. Environ. Hyg.* 15:399–412. doi:10.1080/15459624.2018.1445260.
- Fent, K. W., D. E. Evans, D. Booher, J. D. Pleil, M. A. Stiegel, G. P. Horn, and J. Dalton. 2015. Volatile organic compounds off-gassing from firefighters' personal protective equipment ensembles after use. *J. Occup. Environ. Hyg.* 12:404–14. doi:10.1080/15459624.2015.1025135.
- Fernando, S., L. Shaw, D. Shaw, M. Gallea, L. VandenEnden, R. House, D. K. Verma, P. Britz-McKibbin, and B. E. McCarry. 2016. Evaluation of firefighter exposure to wood smoke during training exercises at burn houses. *Environ. Sci. Technol.* 50:1536–43. doi:10.1021/acs.est.5b04752.
- Filipiak, W., P. Mochalski, A. Filipiak, C. Ager, R. Cumeras, C. E. Davis, A. Agapiou, K. Unterkofler, and J. Troppmair. 2016. A compendium of volatile organic compounds (VOCs) released by human cell lines. *Curr. Med. Chem.* 23:2112–31.
- Geer Wallace, M. A., J. D. Pleil, S. Mentese, K. D. Oliver, D. A. Whitaker, and K. W. Fent. 2017. Calibration and performance of synchronous SIM/scan mode for simultaneous targeted and discovery (non-targeted) analysis of exhaled breath samples from firefighters. *J. Chromatogr. A* 1516:114–24. doi:10.1016/j.chroma.2017.07.082.
- Gómez, M. J., M. M. Gómez-Ramos, A. Agüera, M. Mezcuca, S. Herrera, and A. R. Fernández-Alba. 2009. A new gas chromatography/mass spectrometry method for the simultaneous analysis of target and non-target organic contaminants in waters. *J. Chromatogr. A* 1216:4071–82. doi:10.1016/j.chroma.2009.02.085.
- Hakim, M., Y. Y. Broza, O. Barash, N. Peled, M. Phillips, A. Amann, and H. Haick. 2012. Volatile organic compounds of lung cancer and possible biochemical pathways. *Chem. Rev.* 112:5949–66. doi:10.1021/cr300174a.
- Horn, G. P., R. M. Kesler, S. Kerber, K. W. Fent, T. J. Schroeder, W. S. Scott, P. C. Fehling, B. Fernhall, and D. L. Smith. 2018. Thermal response to firefighting activities in residential structure fires: Impact of job assignment and suppression tactic. *Ergonomics* 61:404–19. doi:10.1080/00140139.2017.1355072.
- Jayatilaka, N. K., P. Restrepo, L. Williams, M. Ospina, L. Valentin-Blasini, and A. M. Calafat. 2017. Quantification of three chlorinated dialkyl phosphates, diphenyl phosphate, 2,3,4,5-tetrabromobenzoic acid, and four other organophosphates in human urine by solid phase extraction-high

- performance liquid chromatography-tandem mass spectrometry. *Anal. Bioanal. Chem.* 409:1323–32. doi:10.1007/s00216-016-0061-4.
- Jo, W.-K., and K.-B. Song. 2001. Exposure to volatile organic compounds for individuals with occupations associated with potential exposure to motor vehicle exhaust and/or gasoline vapor emissions. *Sci. Total Environ.* 269:25–37.
- Keir, J. L. A., U. S. Akhtar, D. M. J. Matschke, T. L. Kirkham, H. M. Chan, P. Ayotte, P. A. White, and J. M. Blais. 2017. Elevated exposures to polycyclic aromatic hydrocarbons and other organic mutagens in Ottawa firefighters participating in emergency, on-shift fire suppression. *Environ. Sci. Technol.* 51:12745–55. doi:10.1021/acs.est.7b02850.
- Kim, D., M. E. Andersen, and L. A. Nylander-French. 2006. Dermal absorption and penetration of jet fuel components in humans. *Toxicol. Lett.* 165:11–21. doi:10.1016/j.toxlet.2006.01.009.
- Kirk, K. M., and M. B. Logan. 2015. Structural fire fighting ensembles: Accumulation and off-gassing of combustion products. *J. Occup. Environ. Hyg.* 12:376–83. doi:10.1080/15459624.2015.1006638.
- Laitinen, J., M. Mäkelä, J. Mikkola, and I. Huttu. 2012. Firefighters' multiple exposure assessments in practice. *Toxicol. Lett.* 213:129–33. doi:10.1016/j.toxlet.2012.06.005.
- Laitinen, J. A., J. Koponen, J. Koikkalainen, and H. Kiviranta. 2014. Firefighters' exposure to perfluoroalkyl acids and 2-butoxyethanol present in firefighting foams. *Toxicol. Lett.* 231:227–32. doi:10.1016/j.toxlet.2014.09.007.
- Lin, C. Y., Å. M. Wheelock, D. Morin, R. M. Baldwin, M. G. Lee, A. Taff, C. Plopper, A. Buckpitt, and A. Rohde. 2009. Toxicity and metabolism of methylnaphthalenes: Comparison with naphthalene and 1-nitronaphthalene. *Toxicol.* 260:16–27. doi:10.1016/j.tox.2009.03.002.
- Lindstrom, A. B., and J. D. Pleil. 1996. Alveolar breath sampling and analysis to assess exposures to methyl tertiary butyl ether (MTBE) during motor vehicle refueling. *J. Air Waste Manage. Assoc.* 46:676–82. doi:10.1080/10473289.1996.10467502.
- McEachran, A. D., K. Mansouri, S. R. Newton, B. E. J. Beverly, J. R. Sobus, and A. J. Williams. 2018. A comparison of three liquid chromatography (LC) retention time prediction models. *Talanta* 182:371–79. doi:10.1016/j.talanta.2018.01.022.
- McEachran, A. D., J. R. Sobus, and A. J. Williams. 2017. Identifying known unknowns using the US EPA's CompTox chemistry dashboard. *Anal. Bioanal. Chem.* 409:1729–35. doi:10.1007/s00216-016-0139-z.
- Meng, C.-K. 2005. *Improving productivity with synchronous SIM/scan*. Agilent Technologies Publication:5989–3108.
- Milman, B. L., and I. K. Zhurkovich. 2017. The chemical space for non-target analysis. *TrAC Trends Anal. Chem.* 97:179–87. doi:10.1016/j.trac.2017.09.013.
- Miranda, A. I., V. Martins, P. Cascão, J. H. Amorim, J. Valente, C. Borrego, A. J. Ferreira, C. R. Cordeiro, D. X. Viegas, and R. Ottmar. 2012. Wildland smoke exposure values and exhaled breath indicators in firefighters. *J. Toxicol. Environ. Health A* 75:831–43. doi:10.1080/15287394.2012.690686.
- Navarro, K. M., R. Cisneros, E. M. Noth, J. R. Balmes, and S. K. Hammond. 2017. Occupational exposure to polycyclic aromatic hydrocarbon of wildland firefighters at prescribed and wildland fires. *Environ. Sci. Technol.* 51:6461–69. doi:10.1021/acs.est.7b00950.
- Newton, S. R., R. L. McMahan, J. R. Sobus, K. Mansouri, A. J. Williams, A. D. McEachran, and M. J. Strynar. 2018. Suspect screening and non-targeted analysis of drinking water using point-of-use filters. *Environ. Pollut.* 234:297–306. doi:10.1016/j.envpol.2017.11.033.
- O'hara, M. E., T. H. Clutton-Brock, S. Green, and C. A. Mayhew. 2009. Endogenous volatile organic compounds in breath and blood of healthy volunteers: Examining breath analysis as a surrogate for blood measurements. *J. Breath Res.* 3:027005. doi:10.1088/1752-7155/3/2/027005.
- Organtini, K. L., A. L. Myers, K. J. Jobst, E. J. Reiner, B. Ross, A. Ladak, L. Mullin, D. Stevens, and F. L. Dorman. 2015. Quantitative analysis of mixed halogen dioxins and furans in fire debris utilizing atmospheric pressure ionization gas chromatography-triple quadrupole mass spectrometry. *Anal. Chem.* 87:10368–77. doi:10.1021/acs.analchem.5b02463.
- Phillips, M., J. Herrera, S. Krishnan, M. Zain, J. Greenberg, and R. N. Cataneo. 1999. Variation in volatile organic compounds in the breath of normal humans. *J. Chromatogr. B Biomed. Sci. Appl.* 729:75–88.
- Pleil, J. D., and M. A. Stiegel. 2013. Evolution of environmental exposure science: Using breath-borne biomarkers for “discovery” of the human exposome. *Anal. Chem.* 85:9984–90. doi:10.1021/ac402306f.
- Pleil, J. D., M. A. Stiegel, and K. W. Fent. 2014. Exploratory breath analyses for assessing toxic dermal exposures of firefighters during suppression of structural burns. *J. Breath Res.* 8:037107. doi:10.1088/1752-7155/8/3/037107.
- Rager, J. E., M. J. Strynar, S. Liang, R. L. McMahan, A. M. Richard, C. M. Grulke, J. F. Wambaugh, K. K. Isaacs, R. Judson, and A. J. Williams. 2016. Linking high resolution mass spectrometry data with exposure and toxicity forecasts to advance high-throughput environmental monitoring. *Environ. Int.* 88:269–80. doi:10.1016/j.envint.2015.12.008.
- Shimoda, K., and H. Hamada. 2010. Bioremediation of fluorophenols by glycosylation with immobilized marine microalga. *Amphidinium Crassum. Environ. Health Insights* 4:87–91.
- Sobus, J. R., J. F. Wambaugh, K. K. Isaacs, A. J. Williams, A. D. McEachran, A. M. Richard, C. M. Grulke, E. M. Ulrich, J. E. Rager, and M. J. Strynar. 2017. Integrating tools for non-targeted analysis research and chemical safety evaluations at the US EPA. *J. Expo. Sci. Environ. Epidemiol.* 28:411–26. doi:10.1038/s41370-017-0012-y.
- Stec, A. A., K. E. Dickens, M. Salden, F. E. Hewitt, D. P. Watts, P. E. Houldsworth, and F. L. Martin. 2018. Occupational exposure to polycyclic aromatic

- hydrocarbons and elevated cancer incidence in firefighters. *Sci. Rep.* 8:2476. doi:[10.1038/s41598-018-20616-6](https://doi.org/10.1038/s41598-018-20616-6).
- Strategies for Developing Optimal Synchronous SIM-Scan Acquisition Methods-AutoSIM/Scan Setup and Rapid SIM. 2006. *Agilent Technologies, Publication* (5989-5669EN).
- Sun, X., K. Shao, and T. Wang. 2016. Detection of volatile organic compounds (VOCs) from exhaled breath as non-invasive methods for cancer diagnosis. *Anal. Bioanal. Chem.* 408:2759-80. doi:[10.1007/s00216-015-9200-6](https://doi.org/10.1007/s00216-015-9200-6).
- Van Thuyne, W., P. Van Eenoo, and F. T. Delbeke. 2008. Implementation of gas chromatography combined with simultaneously selected ion monitoring and full scan mass spectrometry in doping analysis. *J. Chromatogr. A* 1210:193-202. doi:[10.1016/j.chroma.2008.09.049](https://doi.org/10.1016/j.chroma.2008.09.049).
- Veenaas, C., A. Linusson, and P. Haglund. 2018. Retention-time prediction in comprehensive two-dimensional gas chromatography to aid identification of unknown contaminants. *Anal. Bioanal. Chem.* 410:7931-41. doi:[10.1007/s00216-018-1415-x](https://doi.org/10.1007/s00216-018-1415-x).
- Wang, B., S. Yang, G. Chen, Y. Wu, Y. Hou, and G. Xu. 2008. Simultaneous determination of non-volatile, semi-volatile, and volatile organic acids in tobacco by SIM-Scan mode GC-MS. *J. Sep. Sci.* 31:721-26. doi:[10.1002/jssc.200700318](https://doi.org/10.1002/jssc.200700318).
- Wang, H., M. J. Muehlbauer, S. K. O'Neal, C. B. Newgard, E. R. Hauser, J. R. Bain, and S. H. Shah. 2017. Recommendations for improving identification and quantification in non-targeted, GC-MS-based metabolomic profiling of human plasma. *Metabolites* 7:45. doi:[10.3390/metabo7030045](https://doi.org/10.3390/metabo7030045).