Domestic groundwater wells in Appalachia show evidence of low-dose, complex mixtures of legacy pollutants

Nicolette A. Bugher[1], Boya Xiong[1,4], Runako I. Gentles[1], Lukas D. Glist[1], Helen G. Siegel[2], Nicholaus P. Johnson[3], Cassandra J. Clark[3], Nicole C. Deziel[3], James E. Saiers[2], Desiree L. Plata[1]

[1]Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, Parsons Laboratory, 15 Vassar Street, Cambridge Massachusetts 02139 USA

[2]Yale School of the Environment, Environmental Science Center, 21 Sachem Street, New Haven, Connecticut 06511, USA

[3]Yale School of Public Health, Department of Environmental Health Sciences, 60 College St., New Haven, Connecticut 06510

[4]University of Minnesota, Department of Civil, Environmental, and Geo-Engineering, 500 Pillsbury Drive S.E., Minneapolis, MN 55455

Supplemental Information

Table S1: VOC compound index, limit of detection (LOD), and applicable maximum contamination levels (MCLs). *Note:* reported detection limits were established from the average of daily individual LODs.

Index	Compound	LOD (ug/L)	EPA MCL (ug/L)
1	chloromethane	4x10^{-3}	-
2	vinyl chloride	4x10^{-5}	0.002
3	bromomethane	1x10^{-3}	-
4	chloroethane	3x10^{-3}	-
5	trichlorofluoromethane	2x10^{-3}	-
6	1,1 dichloroethylene and trans-1,2-dichloroethylene	3x10^{-3}	0.007 and 0.1
7	1,1-dichloroethane	1x10^{-2}	-
8	2,2 dichloropropane and cis-1,2-dichloroethylene	2x10^{-3}	0.07
9	bromochloromethane	2x10^{-2}	-
10	chloroform	6x10^{-3}	0.08[a]
11	1,1,1-trichloroethane	1x10^{-2}	0.2
12	1,1-dichloropropene and carbon tetrachloride	2x10^{-2}	0.005

13	1, 2-dichloroethane and benzene	4x10^{-4}	0.005 and 0.005
14	trichloroethene	8x10^{-3}	0.005
15	1,2-dichloropropane	2x10^{-3}	0.005
16	dibromomethane	6x10^{-3}	-
17	bromodichloromethane	2x10^{-2}	0.08[a]
18	cis-1,3-dichloropropene	1x10^{0}	-
19	toluene	1x10^{-4}	1
20	trans-1,3-dichloropropene	5x10^{-3}	-
21	1,1,2-trichloroethane	8x10^{-2}	0.005
22	1,3-dichloropropane	7x10^{-3}	-
23	tetrachloroethene	1x10^{-2}	0.005
24	dibromochloromethane	6x10^{-2}	0.08[a]
25	1,2-dibromoethane (EDB)	5x10^{-2}	-
26	chlorobenzene	1x10^{-3}	0.1
27	1,1,1,2-tetrachloroethane	6x10^{-3}	-
28	ethylbenzene	2x10^{-3}	0.7
29	m-xylene and p-xylene	1x10^{-3}	10[b]
30	o-xylene, styrene, bromoform	2x10^{-3}	10[b] and 0.08[a]
31	1,1,2,2-tetrachloroethane	9x10^{-2}	-
32	isopropylbenzene (cumene) and n-propylbenzene	6x10^{-3}	-
33	bromobenzene	9x10^{-3}	-
34	1,2,3-trichloropropane	4x10^{-2}	-
35	1,3,5-trimethylbenzene	8x10^{-3}	-
36	2-chlorotoluene	5x10^{-2}	-
37	4-chlorotoluene	2x10^{-3}	-
38	tert-butylbenzene	4x10^{-3}	-
39	1,2,4-trimethylbenzene	5x10^{-3}	-
40	sec-butylbenzene	5x10^{-2}	-
41	1,3-dichlorobenzene	5x10^{-2}	-
42	p-isopropyltoluene (p-cymene)	1x10^{-2}	-
43	1,4-dichlorobenzene	4x10^{-3}	0.075
44	n-butylbenzene and 1,2-dichlorobenzene	4x10^{-4}	N/A and 0.6
45	1,2-dibromo-3-chloropropane	1x10^{-1}	0.0002
46	1,2,4-trichlorobenzene	9x10^{-2}	0.07
47	hexachlorobutadiene	2x10^{-2}	-
48	naphthalene	9x10^{-3}	-
49	1,2,3-trichlorobenzene	1x10^{-2}	-
50	methylene chloride (dichloromethane)	5x10^{-3}	0.005

[a]MCL for Total Trihalomethanes (sum of 4). [b]MCL for xylenes (Total).

Quality Assurance and Quality Control Assessment of study quality involved interrogation of blank control samples (i.e., field blanks and laboratory blanks) to assess potential procedural or analytical contamination. VOC detection frequencies were less than 10% for all compounds except dichloromethane in laboratory blanks (46%), which was interrogated further for potential laboratory contamination bias. VOCs were detected above limits of detection (**SI Table 1**) in field blanks in less than 10% of samples for all but five compounds (chloromethane; 16%, vinyl chloride; 14%, bromomethane; 15%, dichloromethane; 53%, bromochloromethane; 29%) (**SI Appendix I Data file**). Field blank analyte concentrations were not normally distributed. Thus, to investigate bias, non-parametric distribution-free statistical analyses were conducted for each of the designated VOCs as outlined in Hahn and Meeker (1991) utilizing ranked sample concentrations and binomial probability to determine an upper confidence limit (UCL) on VOCs with a high frequency and magnitude of field blank detections. This blank UCL allows for inference on potential contamination in the population of environmental samples; i.e., the 90th-percentile concentration of blank samples for those with frequent VOC detections is the anticipated maximum contamination based on USGS quality assurance for VOCs (Bender et al., 2011). Dichloromethane UCL calculated LOD was 8.13 ug/L, which far exceeded the analytically determined LOD of 0.005 ug/L. This level was similar to the DCM concentration distribution in environmental samples, and routinely appeared in laboratory blanks. Thus, contamination bias was *not* disproven, and dichloromethane was removed from the analysis accordingly. Vinyl chloride (OH and WV UCL adjusted LOD: 0.06 ug/L), chloromethane (WV UCL adjusted LOD: 0.08 ug/L), bromomethane (OH UCL adjusted LOD: 0.1 ug/L) and bromochloromethane (OH and WV UCL adjusted LOD: 0.4 ug/L) were not detected in laboratory blanks and considered free of laboratory contamination. Accordingly, we included vinyl chloride, chloromethane, bromomethane and bromochloromethane detections at reported LODs **(SI Table 1**) to accurately reflect concentrations in the water, as laboratory blanks were free of contamination.

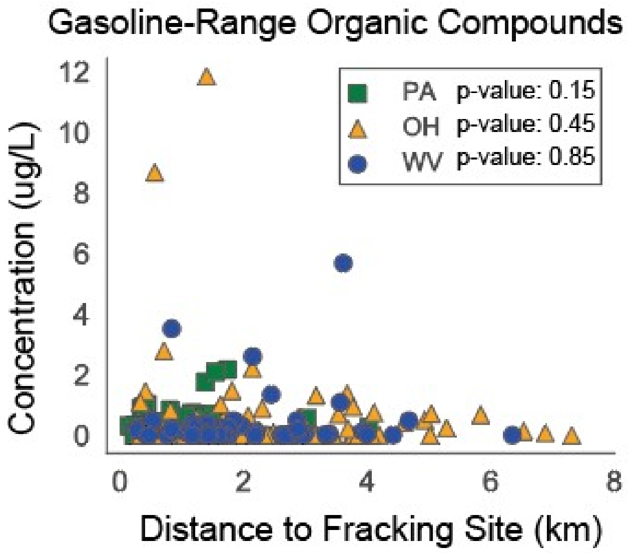

Figure SI 1: Correlation analyses of distance to hydraulic fracturing well and gasoline-range organic compound concentrations (ug/L) for samples taken in West Virginia, Ohio and Pennsylvania.

Gasoline- and Diesel Range Organic Compounds as Mixture Screening Tools. As a non-specific measure of total very volatile organic compound exposure, this study measured gasoline range organic compound concentrations (sum of compounds eluting between characteristic compounds C6 to C10) for each sample. Low levels of GRO concentrations were detected in samples from Ohio and West Virginia, 0.04 ± 0.05 to 11.8 ± 0.2 ug/L and 0.01 ± 0.01 to 5.6 ± 0.8 ug/L, respectively. (Note that PA GRO levels from this sampling effort were 0.13 ± 0.06 to 2.2 ± 0.7 ug/L, as reported in Xiong et al. 2022[1]). GRO concentrations were not correlated with distance to closest hydraulic fracturing well (Spearman's Correlation; p-value > 0.05) (**Figure SI 1**). GRO compound measurements are integrated measures to capture mixture tendencies and occurrences in water samples. Samples with high GRO compound concentrations can be interrogated to determine specific chemical constituents. Here, we see GRO compound concentrations linearly correlated ($R^2 = 0.83$) to the sum of total VOC (tVOC) concentrations in this study (n = 50). GRO compound concentrations had no linear relationship ($R^2 = 0.172$) to number of chemicals detected per sample site in our target analysis. Since chemical mixture risk assessments rely on both concentration of individual mixture components and number of components in a given sample, both total VOC detections and individual concentrations play a vital role. This study is limited based on the semi-target analysis of only 59 potential VOCs detected in HDHF activities and non-target assessments characterizing all GRO VOC components would strengthen mixture risk assessments.

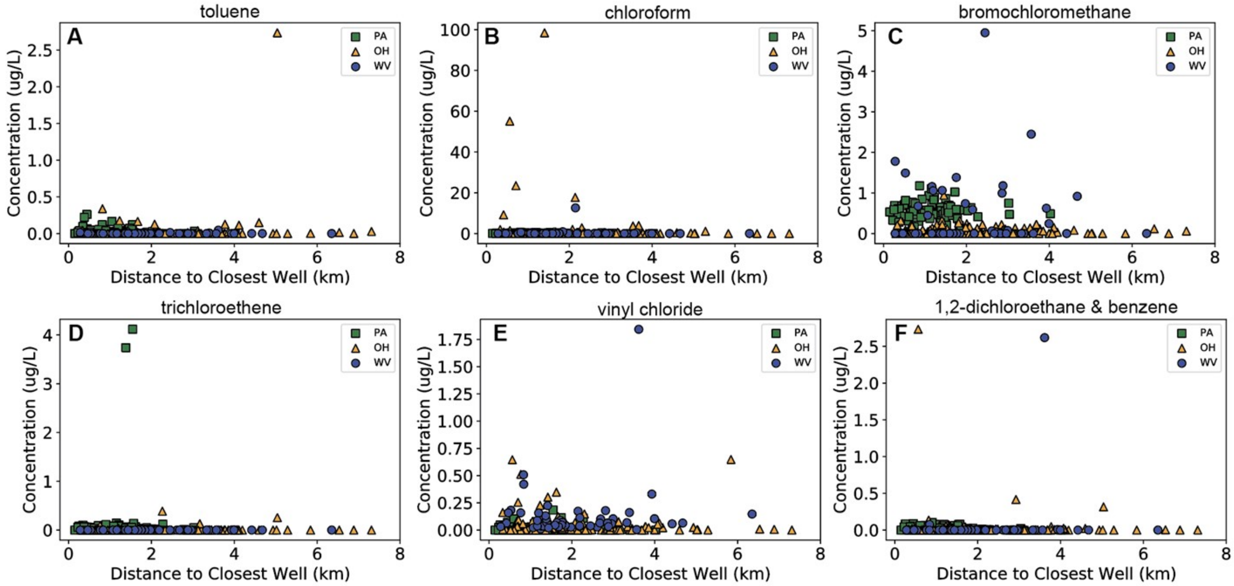

Figure SI 2) Distance to hydraulic fracturing well and concentrations of commonly occurring VOCs: (a) toluene, (b) chloroform, (c) bromochloromethane, (d) trichloroethene, (e) vinyl chloride, and (f) the sum of 1,2-dichloroethane & benzene).

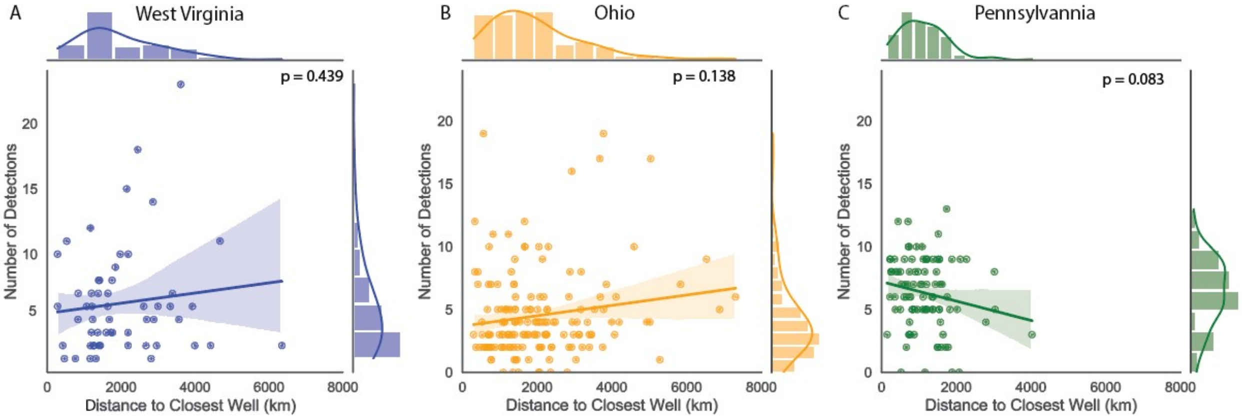

Figure SI 3) Distance to hydraulic fracturing well and number of VOC detections per sampling site. The linear distance between the sample site location and nearest hydraulic fracturing well for (a) West Virginia (histogram range: < 300m to 6.4 km); (b) Ohio (histogram range: < 300 to 7.4 km), and (c) Pennsylvania (histogram range: < 200m to 4.2km) are shown in the histograms above each figure against distribution of number of detections for each state.

Spearman correlation coefficients (p-values between 0.72 and 0.95) indicate no significant linear relationship between distance to closest hydraulic fracturing well site and individual VOC concentrations (**Figure SI2**). In Pennsylvania, concentrations of TCE in wells within 10 km had a weak negative correlation ($r^2 = 0.53$; p-value <0.05, Spearman Correlation) with distance to closest well, but were not significantly elevated (p-value >0.05, Mann Whitney U) compared to concentrations in wells more than 10 km away. The many potential sources of TCE in groundwater may obfuscate pathways from specific contamination site. Ohio and West Virginia showed no significant linear relationships between distance and individual VOCs concentrations. Additionally, total VOC concentrations ($VOC_{TOT;}$ data not shown) and number of VOC detections (**Figure SI3**) were not significantly correlated in all three states (Spearman Rank Order Correlation; r^2 between 0.16 and 0.52; p-value > 0.05), suggesting limited applicability of frequency of detections of tVOCs for use as a measure of “human impact” as seen in previous studies[2]. Nevertheless, the variation in detection frequency of singular VOCs between states provides evidence of local variation in water composition between sampling areas.

References

1. Xiong, B., Soriano, M. A., Gutchess, K. M., Hoffman, N., Clark, C. J., Siegel, H. G., Andrew, G., de Vera, D., Li, Y., Brenneis, R. J., Cox, A. J., Ryan, E. C., Sumner, A. J., Deziel, N. C., Saiers, J. E., & Plata, D. L. (2022). Groundwaters in Northeastern Pennsylvania near intense hydraulic fracturing activities exhibit few organic chemical impacts †. Environmental Science and Technology, https://doi.org/10.1039/d1em00124h
2. Deeds, D. A., Kulongoski, J. T., & Belitz, K. (2012). Assessing California Groundwater Susceptibility Using Trace Concentrations of Halogenated Volatile Organic Compounds. Environmental Science and Technology. https://doi.org/10.1021/es303546b
3. Hahn, G.J. and Meeker, W.Q. (1991) Statistical Intervals: A Guide for Practitioners, John Wiley & Sons, New York. http://dx.doi.org/10.1002/9780470316771
4. Bender, D.A., Zogorski, J.S., Mueller, D.K., Rose, D.L., Martin, J.D., and Brenner, C.K., (2011) Quality of volatile organic compound data from groundwater and surface water for the National Water-Quality Assessment Program, October 1996–December 2008: U.S. Geological Survey Scientific Investigations Report 2011–5204, 128 p.