Supporting Information for Assessing unconventional oil and gas exposure in the Appalachian Basin: Comparison of exposure surrogates and residential drinking water measurements

Cassandra J. Clark[1], Boya Xiong[2,3], Mario A. Soriano, Jr.[4], Kristina Gutchess[4], Helen G. Siegel[4], Emma C. Ryan[5], Nicholaus P. Johnson[1], Kelsie Cassell[6], Elise G. Elliott[1,7,8], Yunpo Li[2], Austin J. Cox[2], Nicolette Bugher[2], Lukas Glist[2], Rebecca J. Brenneis[2], Keli M. Sorrentino[9], Julie Plano[9], Xiaomei Ma[10], Joshua L. Warren[11], Desiree L. Plata[2], James E. Saiers[4], and Nicole C. Deziel[1]

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

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

[3] University of Minnesota, Department of Civil, Environmental and Geo-Engineering, 500 Pillsbury Dr. SE, Minneapolis, MN 55455, USA

[4] Yale School of the Environment, 195 Prospect Street, New Haven, Connecticut 06511, USA

[5] Tufts University, Department of Public Health and Community Medicine, 136 Harrison Avenue, Boston, MA 02111, USA

[6] Yale School of Public Health, Department of Epidemiology of Microbial Diseases, 60 College St., New Haven, Connecticut 06510, USA

[7] Harvard T.H. Chan School of Public Health, Department of Environmental Health, Boston, MA 02115, USA

[8] Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA

[9] Center for Perinatal, Pediatric and Environmental Epidemiology, Yale University Schools of Public Health and Medicine, 1 Church Street, New Haven, CT, 06510, USA

[10] Yale School of Public Health, Department of Chronic Disease Epidemiology, 60 College St., New Haven, Connecticut 06510, USA

[11] Yale School of Public Health, Department of Biostatistics, 60 College St., New Haven, Connecticut 06510, USA

Number of pages: 9
Number of tables: 2
Number of figures: 1

Study Zip Code Selection and Participant Recruitment Criteria

The study areas of Bradford County, Pennsylvania (PA) and Belmont and Monroe Counties, Ohio (OH) were selected based upon several criteria, including high number of unconventional oil and gas wells, low number of conventional wells, high proportion of domestic water well users, and prior oil and gas contamination events (e.g., [1]). We considered inclusion of Class II injection wells in our proximity metrics in addition to UOG wells, but there were none in our PA counties, only two in Belmont County, and zero in Monroe County. We used multiple outreach methods to recruit participants with a range of backgrounds and characteristics representative of the study area. Postcards with information about the study, the study phone number, and the study website were distributed by mail to every address within each eligible zip code. Flyers were posted at local businesses, and virtual ads were run on social media. Interested participants who responded to our recruitment methods were screened for eligibility via phone by study staff, and if eligible, scheduled for a home visit. Study eligibility consisted of being an adult household decision-maker (≥21 years of age), able to communicate in English, and living in our selected counties in a home served by a private groundwater well or spring. The study protocol was approved by the Institutional Review Board of Yale University (HIC #2000021809) and reviewed and approved by the US Environmental Protection Agency (HSR-001162). All participants provided informed consent prior to data collection activities.

Water Sampling and Analytical Methods

Water samples were generally collected from outdoor spigots or hand pumps**.** Prior to sampling, the well or spring was purged until temperature, pH, conductivity, and dissolved oxygen were stable. Stability of the groundwater was assessed using either a YSI 556 Handheld Multiparameter Instrument or YSI Professional Plus with a flow through cell.

With respect to organic compounds, water samples analyzed for volatile organic compounds (VOC) were collected in precombusted clear glass vials containing 1 mL of 50% (v/v) hydrochloric acid (final pH <2) with less than 0.5 mL of headspace and stored on ice at 4°C prior to analysis within four weeks. For analysis of VOCs, we followed U.S. Environmental Protection Agency (U.S. EPA) Method 624, with minor modifications previously described by Getzinger et al.[2] Samples were collected and analyzed in triplicate, and the final reported value is the average of three samples. For the VOCs, limit of detection (LOD) and limit of quantification (LOQ) values were calculated for each compound. Values below the LOD were considered not detected, while measurements above the LOD including those between the LOD and LOQ were retained and reported. Values between the LOD and LOQ were calculated based on an extrapolation from the calibration curve. Chemicals were reported together as groups when their respective chromatographic peaks could not be differentiated.

With respect to inorganic elements and compounds, water samples analyzed for major cations, major anions, and trace elements were filtered in-line through 0.45 µm polyethersulfone filter membranes, collected in 125-mL acid-washed HDPE bottles, and stored at 4°C (major cations and trace elements) or frozen (major anions) until analysis.[3] Samples collected for trace element and major cation analyses were preserved with 50% v/v trace metal grade HNO_3 to a pH of less than 2.0. Major cations and dissolved iron concentrations were quantified by inductively coupled plasma emission spectrometry at the Cary Institute for Ecosystem Studies. Concentrations of major anions and remaining trace elements were determined at the Yale Analytical and Stable Isotope Center using ion chromatography and inductively coupled plasma mass spectrometry, respectively. The LOD for each chemical was defined using the standard deviation of measurements from field blank samples.

Table S1. Organic and inorganic compounds included in analysis.

Compound Name
Organic chemicals
1,2-Dichloroethane and Benzene
1,1-Dichloroethene and trans-1,2-Dichloroethene
1,1,1,2-Tetrachloroethane
1,1,1-Trichloroethane
1,1,2,2-Tetrachloroethane
1,1,2-Trichloroethane
1,1-Dichloroethane
1,1-Dichloropropene and Carbon Tetrachloride
1,2,3-Trichlorobenzene
1,2,3-Trichloropropane
1,2,4-Trichlorobenzene
1,2,4-Trimethylbenzene
1,2-Dibromo-3-chloropropane
1,2-Dibromoethane
1,2-Dichloropropane
1,3,5-Trimethylbenzene
1,3-Dichlorobenzene
1,3-Dichloropropane
1,4-Dichlorobenzene
2,2-Dichloropropane and cis-1,2-Dichloroethene
2-Chlorotoluene
4-Chlorotoluene
Bromobenzene
Bromochloromethane
Bromodichloromethane
Bromomethane
Chlorobenzene
Chloroethane
Chloroform

Chloromethane
cis-1,3-Dichloropropene
Dibromochloromethane
Dibromomethane
Ethylbenzene
Hexachlorobutadiene
Isopropylbenzene (cumene) and n-Propylbenzene
Methylene Chloride (Dichloromethane)
m-Xylene and p-Xylene
Naphthalene
n-Butylbenzene and 1,2-Dichlorobenzene
o-Xylene, Styrene, Bromoform
p-Isopropyltoluene (p-Cymene)
sec-Butylbenzene
tert-Butylbenzene
Tetrachloroethene
Toluene
trans-1,3-Dichloropropene
Trichloroethene
Trichlorofluoromethane
Vinyl chloride
Inorganic chemicals
Arsenic
Barium
Bromide
Calcium
Chloride
Fluoride
Iron
Potassium
Lithium

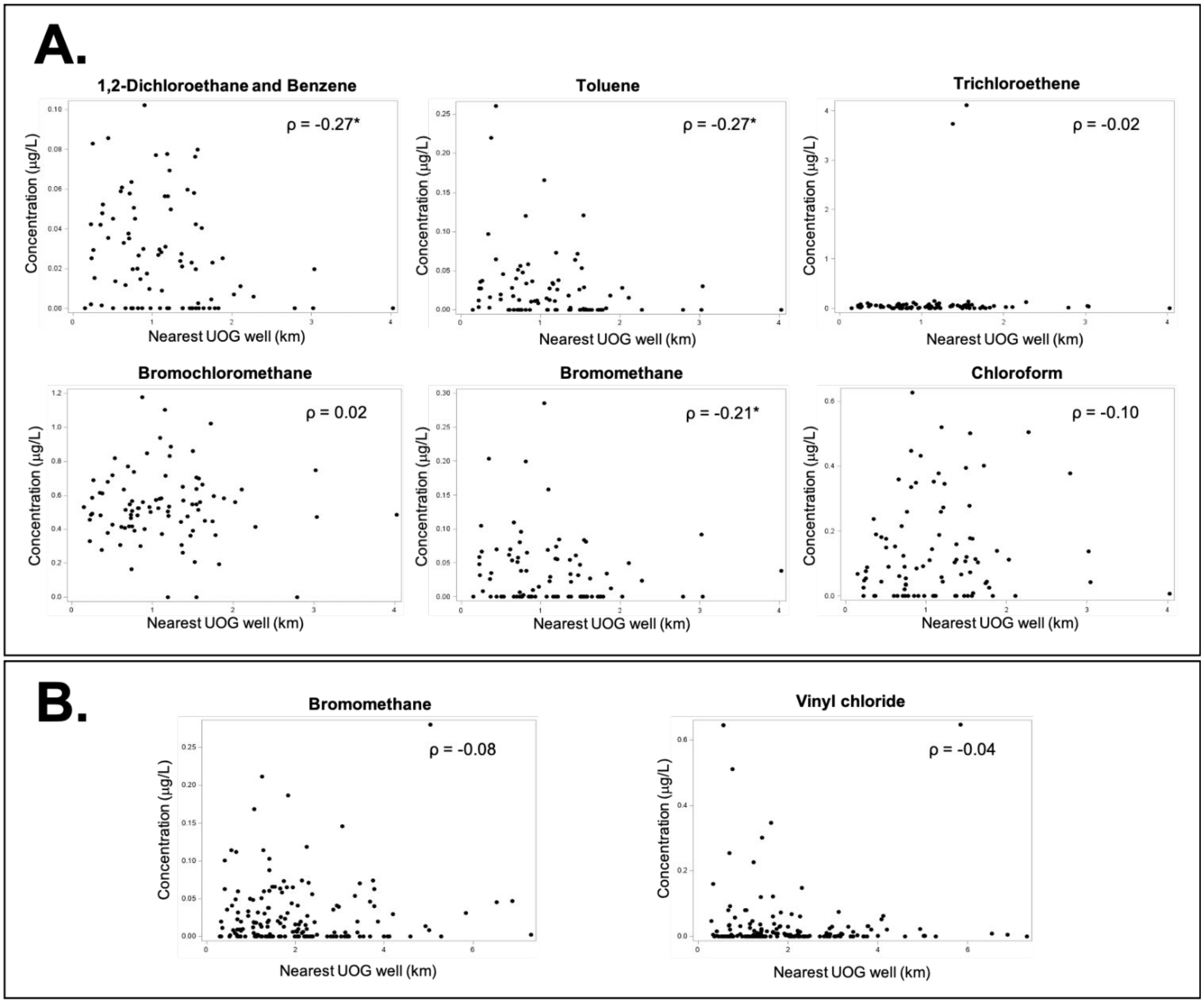
Magnesium
Manganese
Sodium
Nitrate
Lead
Sulfate
Strontium
Uranium

Table S2. Spearman correlation coefficients between concentrations of inorganic solutes and spatial metrics.

	Nearest UOG well	IDW 2 km	ID^{2}W 2 km	IDW 5 km	ID^{2}W 5 km	ID$_{ups}$ 0.5 km	ID$_{ups}$ 1 km	ID$_{ups}$ 2 km	Sum 2 km	Sum 5 km
Chemical										
PA										
Arsenic	-0.01	-0.02	-0.03	0.06	0.00	0.05	0.08	0.07	0.03	0.09
Barium	-0.15	0.10	0.13	0.15	0.15	0.14	0.17*	0.12	0.09	0.14
Bromide	-0.09	0.16*	0.16*	0.06	0.11	0.27*	0.22*	0.17*	0.14	0.05
Calcium	0.07	-0.10	-0.11	-0.18*	-0.17*	-0.10	-0.05	-0.06	-0.05	-0.21*
Chloride	-0.25*	0.23*	0.24*	0.07	0.12	0.27*	0.24*	0.22*	0.23*	0.02
Fluoride	0.07	-0.06	-0.07	-0.04	-0.05	0.03	-0.05	-0.05	-0.08	0.00
Iron	0.12	-0.19*	-0.16*	-0.26*	-0.20*	-0.03	-0.05	-0.07	-0.21*	-0.28*
Lead	0.00	0.00	0.00	0.05	0.05	-0.02	0.06	0.12	-0.01	0.02
Lithium	0.02	0.00	0.00	-0.04	-0.04	-0.03	-0.03	0.00	0.04	-0.04
Magnesium	0.18*	-0.16*	-0.18*	-0.25*	-0.24*	-0.15	-0.14	-0.15	-0.08	-0.26*
Manganese	0.22*	-0.15	-0.15	-0.16*	-0.15	-0.04	-0.10	-0.11	-0.12	-0.15
Nitrate	-0.11	0.02	0.02	0.13	0.08	0.01	0.07	0.11	0.02	0.16*
Sodium	0.03	0.01	0.00	-0.03	-0.03	0.06	0.04	0.06	0.05	-0.03
Strontium	0.01	0.01	0.01	-0.11	-0.07	-0.01	0.01	-0.01	0.04	-0.14
Sulfate	0.13	-0.13	-0.16*	-0.23*	-0.22*	-0.17*	-0.12	-0.12	-0.06	-0.24*
Uranium	-0.05	-0.01	0.00	0.13	0.08	0.03	0.11	0.16*	-0.01	0.15
OH										
Barium	0.09	-0.01	-0.02	-0.02	-0.05	0.06	0.00	-0.06	-0.01	-0.02
Bromide	0.09	-0.06	-0.07	-0.17*	-0.15	0.21*	0.03	-0.06	-0.06	-0.15
Calcium	-0.07	0.06	0.07	0.06	0.07	-0.02	0.04	0.01	0.06	0.09
Chloride	0.06	-0.05	-0.07	0.02	-0.03	0.11	0.05	-0.06	-0.04	0.10
Fluoride	0.02	-0.01	-0.01	-0.11	-0.08	0.22*	0.06	0.02	-0.02	-0.11
Iron	-0.06	0.05	0.05	0.07	0.07	0.19	0.16*	0.09	0.03	0.08
Lithium	0.00	-0.02	-0.01	-0.07	-0.03	0.04	-0.03	0.02	-0.02	-0.05
Magnesium	0.00	-0.03	-0.02	0.00	0.00	-0.09	-0.04	-0.03	-0.02	0.08
Manganese	-0.07	0.05	0.06	-0.07	0.00	0.16*	0.10	0.07	0.04	-0.10
Nitrate	-0.02	0.02	0.01	0.10	0.08	-0.16*	-0.04	0.03	0.02	0.06
Sodium	0.07	-0.11	-0.10	-0.20*	-0.16*	0.10	-0.03	-0.03	-0.13	-0.18*
Strontium	0.01	-0.04	-0.03	-0.08	-0.06	-0.01	-0.08	0.01	-0.03	-0.02
Sulfate	-0.15	0.12	0.11	0.11	0.13	-0.01	0.03	0.11	0.12	0.13

* Statistically significant at the $p < 0.05$ level.

Figure S1. Correlations between concentrations of organic chemicals and distance to nearest UOG well in PA (panel A) and OH (panel B).



Scatterplots presented for organic compounds detected in at least 50% of samples. Statistically significant ($p < 0.05$) correlations are denoted with an *asterisk.

References

1. Llewellyn, G. T.; Dorman, F.; Westland, J. L.; Yoxtheimer, D.; Grieve, P.; Sowers, T.; Humston-Fulmer, E.; Brantley, S. L., Evaluating a groundwater supply contamination incident attributed to Marcellus Shale gas development. **2015,** *112*, (20), 6325-6330.
2. Getzinger, G. J.; O'Connor, M. P.; Hoelzer, K.; Drollette, B. D.; Karatum, O.; Deshusses, M. A.; Ferguson, P. L.; Elsner, M.; Plata, D. L., Natural Gas Residual Fluids: Sources, Endpoints, and Organic Chemical Composition after Centralized Waste Treatment in Pennsylvania. *Environmental Science & Technology* **2015,** *49*, (14), 8347-8355.
3. Wilde, F. D. *Water-quality sampling by the U.S. Geological Survey-Standard protocols and procedures*; 2010-3121; 2010.