HEI Energy Research Report 233

Hydrocarbon Extraction and Risk to Groundwater in Pennsylvania: Part 1. Using Geoscientific Analysis and Community Engagement to Analyze Exposures to Potential Groundwater Contamination Related to Hydrocarbon Extraction in Southwestern Pennsylvania

Baka, Brantley, and Xue et al.

Appendices A–B and Additional Materials 1–3

Correspondence may be addressed to Dr. Jennifer Baka, 320 Walker Building, University Park, PA 16802; email: *jeb525@psu.edu.*

Although this report was produced with partial funding by the United States Environmental Protection Agency under Contract No. 68HERC19D0010 to the Health Effects Institute, it has not been subjected to the Agency's peer and administrative review and may not reflect the views of the Agency; thus, no official endorsement by the Agency should be inferred. This report also has not been reviewed by private party institutions, including those that support HEI Energy, and may not reflect the views or policies of these parties; thus, no endorsement by them should be inferred.

Supplementary Appendices A–B and Additional Materials 1–3

Contents

Appendix A: Supporting Tables for the Report

This appendix provides additional details on: descriptive characteristics of the water sampling database (A1); our tests to evaluate the independence of samples (A2-A3); additional statistical tests conducted to supplement analysis in the report (A4-A10).

Additional Details About Collection and Quality Controls for the Water Sample Dataset

Data collection and quality controls for the Shale Network dataset were previously described and published.[35,36] Quoting from one of those publications,[35] "Water analyses were completed by commercial laboratories (e.g., ALS Environmental … Middleton, PA; Test America … Nashville, TN, and Pittsburgh, PA). These laboratories are accredited by the National Environmental Laboratory Accreditation Program (NELAP) and the Department of Defense Environmental Laboratory Approval Program (DoD ELAP), as well as other state and local accreditation programs. The sampling and analyses were contracted by oil and gas companies … to assess water quality before drilling new shale gas wells. Both the consultants conducting the sampling and the commercial laboratories are selected for their expertise in groundwater sampling and analysis. Expertise is important because the gas companies use this data to defend themselves in case homeowners subsequently notice water quality issues. Samples were shipped to commercial laboratories under a chain of custody for analysis.

Sampling of domestic water wells is not as standardized a process as sampling monitoring wells.[37 38] Not all PADEP reports provided information on sampling procedures, nor on the exact location where samples were collected. General sample location, purging technique, and field measurements were noted on all laboratory reports. …Samples were most commonly collected from spigots on lines coming from groundwater wells that were associated with pressure tanks. However, occasionally samples were taken from kitchen or bathroom faucets, barn spigots, garage spigots, or directly from a wellhead or spring. According to standard practice, a sampling point must be located upgradient of any water treatment or filter systems. Typically, water is run for 15 minutes to purge before the sample is collected.

All laboratory reports containing data were shared with landowners and then released to the PA DEP. The data have been made publicly accessible through Shale Network.[34] Samples were analyzed using standard EPA-approved methods. The reports provided measurements of field parameters, volatile organic compounds, total inorganic anions, total inorganic cations, turbidity, dissolved gases, specific conductivity, total dissolved solids, total suspended solids, alkalinity, pH, hexane extractable material, and surfactants...

[One of the] …laboratories also provided information on analyses of blanks, control spikes, and matrix spikes as part of their quality and control procedures. Spikes were required to show between 80 - 120% recovery and duplicates within 33% reproducibility. Results were well within these limits.

According to the ALS Environmental quality assurance information, these checks were performed as well, but specific results were not provided in the reports.

Data were organized by sampling site such that each well or spring tested was considered an individual site. If measurements were below the reporting limit, this limit was noted, and the analysis was indicated as "less than" (lt). If no GPS coordinates were provided, sites were assigned the coordinates of the geographic center of the land parcel from which the sample was taken...

… this study focused only on groundwater [samples]. Therefore, statistical analyses were performed on data only from water wells and springs…The only analyses we rejected were measured for a sample

outside of [our target area] or if it lacked geographic information or if it was not an analysis of groundwater."

Independence of Samples

Statistical tests were run to test the independence of the samples. The assumption of independence underlying interpretations of the Bruner-Munzel and Wilcox-Mann tests (Tables A2, A3) was tested for spatial autocorrelation using Moran's I. ArcGIS Pro default settings were used to determine the distance threshold, which meant spatial relationships were conceptualized based on inverse distance, with a distance threshold of 14400m. Although a small degree of spatial autocorrelation was determined (Moran's I using inverse distances for barium equals 0.046 and for strontium equals 0.024), the z-score (Moran's I standard deviation = 23 [Ba], 9 [Sr], and the *P* values are both <0.05. The results were not adjusted for spatial autocorrelation. Future work could seek to explore the effects of spatial autocorrelation.

Table A1: Available Analytes in the Groundwater Dataset for SW PA

Species	Total # of Samples Detecting This Analyte[a]	# Above Reporting Limit[b]	EPA MCL/SMCL (mg/L)	# at or Above EPA MCL/SMCL	Association with UOGD Processes
		Organic analytes			
Methane	7,100	1,014	10	145	Casing/cementing failures
Ethane	7,097	87	-	-	Casing/cementing failures
Propane	5,150	7	-	-	Casing/cementing failures
Benzene	6,752	8	0.005	3	Fracking fluid leaks/spills
Toluene	6,750	30	1	1	Fracking fluid leaks/spills
Ethylbenzene	6,752	8	0.7	1	Fracking fluid leaks/spills
Xylene	5,824	8	10	1	Fracking fluid leaks/spills
Oil and grease	6,944	634	-	-	Fracking fluid leaks/spills
Methylene Blue Active Substances	6,663	369	-	-	Fracking fluid leaks/spills
		Major inorganic analytes			
Alkalinity[c]	7,012	6,952	-	-	Background geochemistry, AMD
Bromide	1,234	256	-	-	Brine spills/leaks
Calcium	7,103	7,003	-	-	Background geochemistry, AMD
Chloride	6,974	5,633	250	108	Brine spills/leaks
Potassium	6,496	6,148	-	-	Background hydrogeo-chemistry

Species	Total # of Samples Detecting This Analyte[a]	# Above Reporting Limit[b]	EPA MCL/SMCL (mg/L)	# at or Above EPA MCL/SMCL	Association with UOGD Processes
Magnesium	7,101	6,916	-	-	Background geochemistry, AMD
Sodium	7,110	7,093	-	-	Brine spills/leaks
Sulfate	7,121	6,845	250	190	Background geochemistry, AMD
Total dissolved solids	7,122	7,111	500	1,019	Background hydrogeo-chemistry
pH	7,085	7,076	-	-	Background hydrogeo-chemistry
Trace inorganic analytes					
Arsenic	6,230	181	0.01	79	Methane migration
Barium	7,096	6,942	2	31	Brine spills/leaks
Iron	7,104	5,326	0.3	2,125	Methane migration
Manganese	7,119	4,144	0.05	2,049	Methane migration
Strontium	6,407	6,265	-	-	Brine spills/leaks

AMD = acid mine drainage; EPA = US Environmental Protection Agency; SW PA = Southwestern Pennsylvania; UOGD = unconventional oil and gas development.

[a] Samples where the laboratory report included the respective analyte, either above or below laboratory reporting limits for concentration.

[b] Samples where the censor code is either nc ("not censored") or gt ("greater than"), indicating the concentration value was at or above the reporting limit.

[c] Analyses reporting either "Alkalinity, total" or "Alkalinity, bicarbonate."

Table A2: Comparison of Median Concentrations in SW PA Samples Across Different Hydrocarbon Land Use Classifications Relative to Control Samples (no Hydrocarbon Extraction Within 1 km)

Species	Land Use	Median (mg/L)[a]	Control Group Median (mg/L)[a, b]	*P* Value (BM)[c]	Confidence[d]
Barium	UOGD	0.140	0.100	8.90E-08	***
	COGD	0.099		6.20E-01	
	UOGD+COGD	0.120		2.37E-10	***
	UOGD+COGD+CM	0.100		1.28E-01	
	UOGD+CM	0.093		9.66E-01	
	COGD+CM	0.098		9.57E-01	
	CM	0.087		8.92E-04	***
Calcium	UOGD	85.4	55.5	2.42E-14	***
	COGD	63.2		1.12E-03	
	UOGD+COGD	73.7		5.11E-14	***
	UOGD+COGD+CM	72.9		0.00E+00	***
	UOGD+CM	82.6		1.14E-04	***
	COGD+CM	72.6		0.00E+00	***
	CM	71.0		5.19E-11	***
Chloride	UOGD	8.0	9.8	1.62E-01	
	COGD	8.9		5.62E-01	
	UOGD+COGD	12.5		3.79E-05	***
	UOGD+COGD+CM	12.0		2.11E-04	***
	UOGD+CM	18.2		1.50E-03	***
	COGD+CM	8.46		1.76E-01	
	CM	10.8		1.12E-01	
Sodium	UOGD	9.2	13.6	3.82E-04	***
	COGD	9.4		8.70E-07	***
	UOGD+COGD	9.7		3.99E-02	*
	UOGD+COGD+CM	13.9		2.67E-01	
	UOGD+CM	16.0		7.04E-01	
	COGD+CM	12.9		9.67E-01	
	CM	13.2		8.79E-01	
Strontium	UOGD	0.37	0.30	1.05E-06	***
	COGD	0.29		2.55E-01	
	UOGD+COGD	0.34		1.55E-05	***

Species	Land Use	Median (mg/L)[a]	Control Group Median (mg/L)[a, b]	*P* Value (BM)[c]	Confidence[d]
	UOGD+COGD+CM	0.43		0.00E+00	***
	UOGD+CM	0.40		2.95E-02	*
	COGD+CM	0.43		0.00E+00	***
	CM	0.37		1.06E-05	***
Sulfate	UOGD	35.1	34.0	6.07E-01	
	COGD	31.0		9.73E-05	***
	UOGD+COGD	32.3		1.44E-05	***
	UOGD+COGD+CM	38.0		6.38E-05	***
	UOGD+CM	42.1		4.14E-03	***
	COGD+CM	38.0		5.02E-07	***
	CM	42.4		1.22E-11	***

CM = coal mining; COGD = conventional oil and gas development; SW PA = Southwestern Pennsylvania; UOGD = unconventional oil and gas development.

[a] Samples with concentration values below the reporting limit were assigned the value of the reporting limit in the analysis. Only samples where all six analytes were analyzed were retained in this analysis.

[b] Control group is defined as samples where the distance to UOG wells, COG wells, and coal mining is ≥ 1 km.

[c] Calculated using a two-sided Brunner-Munzel test. Results using a Wilcoxon–Mann–Whitney test yielded similar *P* values.

[d] * = 95% confidence, ** = 99% confidence, *** = 99.5% confidence.

Table A3: Comparison of Median Barium and Strontium Concentrations (in µg/L) for Samples ≤1 km and >1 km from UOGD Parameters

Species	Region	UOG Parameter	Median (≤1 km)	Median (>1 km)	Difference	Two-Sided *P* Value (BM Test)	One-Sided *P* Value (BM Test)[a]	Confidence[b]
Barium	SW PA	UOG well	110	98	12	1.20E-16	6.01E-17	***
		Spill violation	130	99.1	30.9	1.50E-20	7.49E-21	***
Barium	NE PA	UOG well	130	118	12	7.70E-08	3.85E-08	***
		Spill violation	140	120	20	5.90E-04	2.95E-04	***
Barium	Statewide	UOG well	121	110	11	1.08E-14	5.40E-15	***
		Spill violation	134	111	23	1.66E-13	8.31E-14	***
Strontium	SW PA	UOG well	377	350	27	3.86E-06	1.93E-06	***
		Spill violation	373	359	14	0.126	0.063	
Strontium	NE PA	UOG well	295	270	25	0.0133	6.65E-03	***
		Spill violation	353	272	81	4.37E-04	2.18E-04	***
Strontium	Statewide	UOG well	343.5	311	32.5	2.16E-08	1.08E-08	***
		Spill violation	360	318.5	41.5	7.20E-06	3.60E-06	***

BM = Brunner-Munzel; NE PA = Northeastern Pennsylvania; SW PA = Southwestern Pennsylvania; UOG = unconventional oil and gas; UOGD = unconventional oil and gas development.

[a] Calculated using a one-sided Brunner-Munzel test (are median concentrations are greater within 1 km of UOGD).

[b] For one-sided BM tests: * = 95% confidence, ** = 99% confidence, *** = 99.5% confidence.

Table A4: Relationships Among Barium, Strontium, and the Density (Within 1 km) or Distance of UOG Wells or Well Pad Spills[a]

Region	Species	UOGD Metric	Coefficient	Std. Error	*P* Value	Confidence[b]
SW PA	Barium	UOG well density (1 km)	3.00E-02	3.83E-03	5.78E-15	***
		Spill density (1 km)	7.67E-02	9.57E-03	1.26E-15	***
		UOG well distance	-3.07E-05	4.55E-06	1.72E-11	***
		Spill distance	-9.98E-06	9.68E-07	9.43E-25	***
	Strontium	UOG well density (1 km)	1.18E-02	4.15E-03	4.33E-03	***
		Spill density (1 km)	2.64E-02	1.05E-02	1.19E-02	*
		UOG well distance	-2.91E-05	6.22E-06	2.93E-06	***
		Spill distance	-1.47E-05	1.21E-06	9.29E-34	***
NE PA	Barium	UOG well density (1 km)	1.13E-02	6.04E-03	6.27E-02	
		Spill density (1 km)	1.88E-02	1.46E-02	1.98E-01	
		UOG well distance	-9.42E-05	6.06E-06	3.30E-54	***
		Spill distance	-3.35E-05	2.36E-06	2.27E-45	***
	Strontium	UOG well density (1 km)	9.46E-04	7.05E-03	8.93E-01	
		Spill density (1 km)	2.63E-02	1.66E-02	1.14E-01	
		UOG well distance	-7.89E-05	9.17E-06	8.44E-18	***
		Spill distance	-2.23E-05	3.38E-06	4.92E-11	***
Statewide	Barium	UOG well density (1 km)	1.26E-02	3.72E-03	7.46E-04	***
		Spill density (1 km)	3.74E-02	9.26E-03	5.33E-05	***
		UOG well distance	-6.79E-05	4.12E-06	8.35E-61	***
		Spill distance	-1.84E-05	1.07E-06	1.65E-66	***
	Strontium	UOG well density (1 km)	1.78E-02	4.11E-03	1.47E-05	***
		Spill density (1 km)	4.01E-02	1.02E-02	8.36E-05	***
		UOG well distance	-4.71E-05	5.84E-06	7.15E-16	***
		Spill distance	-9.23E-06	1.38E-06	2.27E-11	***

NE PA = Northeastern Pennsylvania; SW PA = Southwestern Pennsylvania; UOG = unconventional oil and gas.

[a] Separate univariate regressions are run for each species and UOGD metric

[b] * = 95%, ** = 99%, *** = 99.5%.

Table A5: Relationships Among Barium, Strontium, and the Density (within 1 km) or Distance of UOG Wells or Well Pad Spills, Including only Higher Elevation Well Pads[a]

Region	Species	UOGD metric	Coefficient	Std. Error	*P* Value	Confidence[b]
SW PA	Barium	UOG well density (1 km)	3.65E-02	4.39E-03	1.10E-16	***
		Spill density (1 km)	1.04E-01	1.12E-02	2.50E-20	***
		UOG well distance	-7.60E-06	2.23E-06	6.38E-04	***
		Spill distance	-8.22E-06	8.59E-07	1.47E-21	***
	Strontium	UOG well density (1 km)	1.82E-02	4.75E-03	1.26E-04	***
		Spill density (1 km)	2.76E-02	1.22E-02	2.40E-02	*
		UOG well distance	-1.77E-05	3.31E-06	8.34E-08	***
		Spill distance	-1.14E-05	1.09E-06	3.17E-25	***
NE PA	Barium	UOG well density (1 km)	1.99E-02	7.43E-03	7.31E-03	**
		Spill density (1 km)	3.79E-02	1.66E-02	2.29E-02	*
		UOG well distance	-5.40E-05	2.90E-06	1.09E-76	***
		Spill distance	-2.45E-05	1.31E-06	2.66E-77	***
	Strontium	UOG well density (1 km)	9.14E-03	8.70E-03	2.93E-01	
		Spill density (1 km)	4.94E-02	1.90E-02	9.47E-03	**
		UOG well distance	-1.77E-05	3.31E-06	8.34E-08	***
		Spill distance	-2.00E-05	1.91E-06	1.31E-25	***
Statewide	Barium	UOG well density (1 km)	1.90E-02	4.40E-03	1.53E-05	***
		Spill density (1 km)	5.90E-02	1.07E-02	3.89E-08	***
		UOG well distance	-3.39E-05	1.99E-06	1.50E-64	***
		Spill distance	-1.71E-05	8.20E-07	6.09E-96	***
	Strontium	UOG well density (1 km)	2.73E-02	4.85E-03	1.79E-08	***
		Spill density (1 km)	5.30E-02	1.17E-02	6.61E-06	***
		UOG well distance	-3.77E-05	3.06E-06	8.85E-35	***
		Spill distance	-1.14E-05	1.10E-06	3.36E-25	***

NE PA = Northeastern Pennsylvania; SW PA = Southwestern Pennsylvania; UOG = unconventional oil and gas.

[a] Separate univariate regressions are run for each species and UOGD metric

[b] * = 95%, ** = 99%, *** = 99.5%.

Table A6: Fixed-Effects Results Analyzing Relationships Among Barium, Strontium, and UOGD Metrics, Including UOG Wells and Documented Spill-Related Violations[a]

Region	Species	UOGD Metric[b]	Coefficient	Std. Error	*P* Value	Confidence[c]
SW PA	Barium	UOG well density	2.30E-02	3.90E-03	3.76E-09	***
		UOG well distance	-2.15E-05	4.93E-06	1.39E-05	***
		Spill violation density	5.58E-02	9.82E-03	1.42E-08	***
		Spill violation distance	-1.02E-05	1.28E-06	1.60E-15	***
NE PA	Barium	UOG well density	1.97E-02	6.07E-03	1.19E-03	***
		UOG well distance	-1.01E-04	6.14E-06	1.26E-60	***
		Spill violation density	2.51E-02	1.47E-02	8.76E-02	
		Spill violation distance	-3.67E-05	2.51E-06	2.01E-48	***
Statewide	Barium	UOG well density	1.83E-02	3.79E-03	1.45E-06	***
		UOG well distance	-6.61E-05	4.28E-06	1.04E-53	***
		Spill violation density	3.65E-02	9.38E-03	1.02E-04	***
		Spill violation distance	-1.89E-05	1.29E-06	3.59E-48	***
SW PA	Strontium	UOG well density	1.18E-02	4.13E-03	4.41E-03	***
		UOG well distance	-1.48E-05	6.60E-06	2.49E-02	*
		Spill violation density	4.47E-02	1.06E-02	2.36E-05	***
		Spill violation distance	-1.10E-05	1.48E-06	2.15E-12	***
NE PA	Strontium	UOG well density	2.16E-02	7.10E-03	2.37E-03	***
		UOG well distance	-7.28E-05	9.41E-06	1.05E-14	***
		Spill violation density	4.80E-02	1.66E-02	3.91E-03	***
		Spill violation distance	-3.31E-05	3.66E-06	1.78E-19	***
Statewide	Strontium	UOG well density	1.86E-02	4.12E-03	6.55E-06	***
		UOG well distance	-4.81E-05	6.00E-06	1.24E-15	***
		Spill violation density	5.50E-02	1.01E-02	6.03E-08	***
		Spill violation distance	-1.16E-05	1.63E-06	1.23E-12	***

NE PA = Northeastern Pennsylvania; SW PA = Southwestern Pennsylvania; UOG = unconventional oil and gas.
[a] Separate univariate regressions are run for each species and UOGD metric
[b] Calculated for 1 km density.
[c] * = 95%, ** = 99%, *** = 99.5%.

Table A7: Coefficients and *P* Values for "Dummy" Variables Included in Our Fixed-Effects Regression. Values Reported Correspond with Results from Regressions Analyzing the Number of UOG Wells Within 1 km (see Table A6).

Species	Variable	Coefficient	Std. Err.	*P* Value	Significance[a]
Barium	COG well	-0.1040221	0.0224755	3.71E-06	*
	Coal mine	-0.0945249	0.0281998	0.00080353	*
	Anticline	0.18055685	0.0233269	1.03E-14	*
	Stream	0.15521894	0.0165865	8.75E-21	*
	Highway	0.15002705	0.0149276	1.01E-23	*
	Fault	0.04494387	0.0588928	0.44538256	
Strontium	COG well	0.20328888	0.0262643	1.05E-14	*
	Coal mine	0.22616464	0.0311509	4.02E-13	*
	Anticline	-0.0702084	0.0293305	0.01668987	*
	Stream	0.06503936	0.0209636	0.00192207	*
	Highway	0.17396165	0.0188052	2.48E-20	*
	Fault	0.48634209	0.0795974	1.02E-09	*

COG = conventional oil and gas; UOG = unconventional oil and gas.
[a] * = $p < 0.05$.

Table A8: Correlation Coefficients and *P* Values for Regressions of log[Barium] and log[Strontium] Concentrations versus log Waste Production Volumes at UOG Wells Within 1 km

	Barium		***Strontium***	
Region	**Coefficient**	***P* Value**	**Coefficient**	***P* Value**
SW PA	0.0106	4.70E-16	0.00607	2.62E-05
NE PA	0.000754	0.703	0.00300	0.181
Statewide	0.00192	0.120	0.00905	3.32E-11

NE PA = Northeastern Pennsylvania; SW PA = Southwestern Pennsylvania; UOG = unconventional oil and gas.

Table A9: Median [Ba] (in mg/L) Comparison Between Samples Located Near 250- or 500-Gallon Spills ("Spill Group") and Samples >1 km/3 km from a Spill in the Statewide Dataset

Spill Group	Control Group	Median [Ba] (Spill Group)	Median [Ba] (Control)	One-Sided *P* Value (BM Test)[a]	Confidence[b]
≥250-gallon spill within 1 km	No spills within 1 km	0.137	0.111	0.074	
≥250-gallon spill within 1 km	No ≥250-gallon spills within 1 km	0.137	0.112	0.086	
≥250-gallon spill within 3 km	No spills within 3 km	0.131	0.106	2.94E-13	***
≥250-gallon spill within 3 km	No ≥250 gallon spills within 3 km	0.131	0.111	1.41E-8	***
≥500-gallon spill within 1 km	No spills within 1 km	0.140	0.111	0.136	
≥500-gallon spill within 1 km	No ≥500 gallon spills within 1 km	0.140	0.112	0.149	
≥500-gallon spill within 3 km	No spills within 3 km	0.122	0.106	2.58E-05	***
≥500-gallon spill within 3 km	No ≥500 gallon spills within 3 km	0.122	0.112	0.007	**

[a] Calculated using a one-sided BM test evaluating the hypothesis that median concentrations are greater near spills.
[b] * = 95%, ** = 99%, *** = 99.5%.

Table A10: Median [Ba] (in mg/L) Comparison Between SW PA Samples Located Within 1 or 3 km of Impoundments Reprimanded by the PADEP ("Impoundment Group") and SW PA Samples >3 km/1 km from a Reprimanded Impoundment

Impoundment Group	Control Group	Median [Ba] (Impoundment Group)	Median [Ba] (Control Group)	One-Sided *P* Value (BM test)[a]	Confidence[b]
Impoundment within 1 km	No impoundment within 1 km	0.1345	0.100	5.65E-04	***
Impoundment within 1 km	No impoundment within 3 km	0.1345	0.099	3.04E-04	***
Impoundment within 3 km	No impoundment within 1 km	0.123	0.100	2.92E-11	***
Impoundment within 3 km	No impoundment within 3 km	0.123	0.099	4.69E-13	***

PADEP = Pennsylvania Department of Environmental Protection; SW PA = Southwestern Pennsylvania.
[a] Calculated using a one-sided BM test evaluating the hypothesis that median concentrations are greater near impoundments.
[b] * = 95%, ** = 99%, *** = 99.5%.

References

For citations included in this appendix, please refer to the References in the main report, beginning on page 30.

Appendix B: Focus Group Coding Scheme

Table B1: Focus Group Coding Scheme

Focus Group Question(s)	**Research Theme**	**Main Code**	**Subcodes**	**Description**
1, 2	**Observed Changes to Water**	Change	Taste; Color; Smell; General	Description of observed changes to groundwater since the start of UOGD
3, 4	**Contaminants & Pathways**	Substances of Concern	Bentonite clay; Barium; Synthetics (Benzene; Tuluene; Xylene etc.); PFAS; Radiation; Radon; Heavy metals; Strontium; Methane; Salts and brines	UOGD-related contaminants and pathways of contamination of interest to participants
		Above-Ground Mechanisms	Well pad runoff; Impoundment ponds; Well pad spills and flowback	
		Underground Mechanisms	Well casing failure; Methane migration; Brine migration; Well communication issues	
		Off Well Pad	Waste disposal and storage; Compressor station runoff; Water withdrawal	

Focus Group Question(s)	Research Theme	Main Code	Subcodes	Description
5	**Health**	Health Effects	Skin burning; Anxiety and stress; Pregnancy issues; Hormone and growth disruption; Organ failure; Impacts to animals; Occupational health; Cancers	Ways people described UOGD affecting community health
10, 11	**What can be done?**	Data	Data gathering; Transparency; Disclosure	Participant recommendations for improving knowledge on UOGD-water-health relationships
		Communication	Scientists-communities; DEP-communities; Health officials-communities	
		Funding	Government; NGOs; Other	
9	**Obstacles**	Academia	Insufficient public engagement and communication	Topics participants identified as barriers to hindering knowledge about UOGD-water-health relationships
		Government	Bureaucracy; Data gaps/limitations; Regulatory limits; Government-Industry relations	
		Industry	Haliburton Loophole/Fluid content disclosure; Industry "aloofness"; Poor data reporting; Operator carelessness	

Focus Group Question(s)	**Research Theme**	**Main Code**	**Subcodes**	**Description**
		Corruption	Government capture; Academic capture; Media bias	
		Lack of Resources	Funding; Time	
		Outreach/Information Dissemination	N/A	
		Environmental Complexities	Overlapping sources of pollution; Geographic specificities	
		Epistemic Limits	Citizen Science lacks precision; High-proof c; Imprecise language & jargon; Lack of local expertise; Limited data resolution; Long research timescales	
		Other Social Factors	Community silence; politicization of science; lack of education; "Too many fires to put out"	
6, 7, 8	**Knowledge Production and Communication**	Sources of Information	Academic Literature; First- and second-hand experiences; Government sources; Mass sources (books; journalism; social media); NGO sources; Citizen science data; Industry literature	Main sources participants rely on to form their opinions on UOGD and how they communicate their concerns

Focus Group Question(s)	Research Theme	Main Code	Subcodes	Description
		Forms of Communication	Talking to policymakers; Peer-reviewed publication; Public presentations and outreach; Self-published materials; Conferences, workshops, summits; News media	
11	**Other Topics of Concern to Communities**	Other	Cancer clusters; Climate change; The Shell cracker plant; Decline in UOGD in SW PA; Region's industrial history; Broader O&G infrastructure; Plastic waste; Truck traffic; Out migration	Events and phenomena that are related to UOGD in interviewees' eyes that may lack a direct connection to the UOGD-water-health relationship

DEP = Department of Environmental Protection; O&G = oil and gas; PFAS = per- and polyfluoroalkyl substances; SW PA = Southwestern Pennsylvania; UOGD = unconventional oil and gas development.

Additional Materials 1: Consent Form

Consent for Exempt Research
The Pennsylvania State University

Title of Project: Using Geoscientific Analysis and Community Engagement to Analyze Exposures to Potential Groundwater Contamination Related to Hydrocarbon Extraction in Southwestern Pennsylvania

Principal Investigator: Dr. Jennifer Baka, Assistant Professor, Dept. of Geography, PSU, 814-865-9656, jeb525@psu.edu
You are being invited to volunteer to participate in a research study. This summary explains information about this research.

Study Purpose: The purpose of this study is to gather community perceptions on the connections, or lack thereof, amongst energy development, groundwater quality, and public health in Southwestern PA. This information will inform a geoscientific analysis of a water sample database of samples collected in Beaver, Washington, and Greene counties.

You are being invited to participate in this study because of your personal experience and firsthand knowledge of energy development, water quality, and public health in the region. Your views and experiences will enrich our understanding of these impacts and processes.

Study Procedures: Your participation in this research will consist of one or two focus groups composed of residents of your county to discuss your own perceptions of and knowledge about the way fossil fuel extraction affects drinking water and public health in Washington, Greene, or Beaver County. The primary aim of the second focus group will be to discuss the preliminary geoscientific analyses, initial analyses of the first focus group, as well as compile policy recommendations for scientists and policymakers. The focus groups will last approximately 90 minutes each and will be audio-recorded.

Confidentiality: Your participation is voluntary, and you may decide to stop at any time. You do not have to answer any questions that you do not want to answer. Your identity will be kept strictly confidential. Your name will not be used in any publications resulting from the research. Generic names and code numbers will be used to identify subjects on all official documents, final study reports, and to all external parties. Audio recordings will be uploaded to a portable laptop and will be password-protected. Original records will be destroyed after the research is completed. Paper documents will be similarly protected. All computerized files will be password-protected and encrypted.

If you have questions or concerns, you should contact Dr. Baka at 814-865-9656 or Owen Harrington at 774-364-5366. If you have questions regarding your rights as a research subject or concerns regarding your privacy, you may contact the Office for Research Protections at 814-865-1775.

Additional Materials 2: Focus Group Recruitment Flyer

Interested in discussing the links between energy development, groundwater quality, and public health? Participate in upcoming focus groups organized by Penn State researchers!

Overview: A Penn State research team is currently researching these topics in Beaver, Washington, and Greene Counties. We are looking to gather community insights at upcoming focus groups, which will consist of 12–15 community members.

Process: The focus groups will last approximately 90 minutes. We will record and transcribe the focus groups. The recordings and transcriptions will only be shared with our research team.
Your participation is voluntary, and you can stop participating at any time. All responses will also be anonymized.

When and where:

County	Date	Time	Location
Beaver	June 23	6:30–8:00 pm	Penn State Beaver Campus
Washington	July 14	6:30–8:00 pm	Washington Park, Stone Pavilion
Greene	July 15	6:30–8:00 pm	Greene County Fairgrounds, Downstairs Building

Why participate? Your input will inform a geoscientific analysis of a water sampling database collected in the region (7,000 samples). Additionally, the findings will be shared with representatives from various government agencies, including the US Environmental Protection Agency, the PA Environmental Protection Agency, and the PA Department of Health, all while keeping your personal data confidential.

As a thank you, light refreshments will be served, and participants will receive a $15 gas card.

Registration: To register, please fill out this form: https://forms.office.com/r/xb4EuXzfBm. Or, get in touch with Dr. Jennifer Baka, the Principal Investigator, or Owen Harrington, a Penn State Graduate Student in Geography, who will be helping to conduct the focus groups.

Dr. Baka: jeb525@psu.edu, 814-865-9656
Mr. Harrington: ofh5033@psu.edu

Funding sources: The study is funded by the Health Effects Institute–Energy, the National Science Foundation, the Institutes of Energy and Environment, and the Earth and Environmental Systems Institute at Penn State.

For more information, please see the project website: https://www.heienergy.org/research/water-quality/baka

Additional Materials 3: Summer 2022 Focus Group Questions

1. How would you describe the drinking water quality in your community?
2. Has the drinking water quality changed over time?
    a. What specific changes have you observed over time?
        i. Follow up with things like:
            1. Any change in water color?
            2. In smell?
            3. In taste?
3. What factors do you think have caused these changes?
4. Do you think energy extraction (i.e., coal mining, conventional oil and gas development, and unconventional oil and gas development) has impacted water quality?
    a. If yes: Which of these industries?
    b. What information/experiences have led you to believe these things?
5. How do you think water quality relates to public health in your community?
    a. What information/experiences have led you to believe these things?
6. Have you seen any information from government sources — such as the PA Department of Health, PA Department of Environment, the US Environmental Protection Agency — on energy development, water quality, and public health?
    a. Have any of these agencies ever contacted you?
    b. If so, which ones?
    c. Do you check any websites to get this information?
        i. If so, which ones?
7. Of the concerns you have voiced up to this point, have you tried to communicate them?
    a. Have you submitted a complaint to a government agency?
    b. Have you reached out to a non-profit or advocacy group for help?
    c. Have you reached out to a company operating near you for help?
    d. If yes, to whom and how?
    e. If no, why not?
    f. Have any of your concerns been resolved?
8. Are you aware of any community-led water testing happening in your area?
    a. Have you ever participated in this testing? If so, which ones?
    b. What were the outcomes of the testing?
9. What, in your mind, are the biggest barriers to improving water quality and public health in Southwestern PA?
    a. Prompts can include:
        i. Politics
        ii. Industry
        iii. Corruption
        iv. Financing/investment
10. If you had the power to change something related to drinking water and/or public health in your area, what would it be?
11. Is there anything we haven't discussed that you think is important to this topic?