Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to [508 standards](#) due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehp508@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

Supplemental Material

Unconventional Oil and Gas Development Exposure and Risk of Childhood Acute Lymphoblastic Leukemia: A Case–Control Study in Pennsylvania, 2009–2017

Cassandra J. Clark, Nicholaus P. Johnson, Mario Soriano Jr, Joshua L. Warren,
Keli M. Sorrentino, Nina S. Kadan-Lottick, James E. Saiers, Xiaomei Ma, and Nicole C. Deziel

Table of Contents

Intermediate analyses of association and correlation to identify covariates and confounders for model building.

Table S1. Intermediate χ^2 or t-test analyses to identify covariates and confounders for model building.

Covariate	Association with outcome	Association with exposure (ID^{2}W 5km, primary window)	Change in OR of >10%?
Sex*	0.57	0.41	N
Gestational age	0.76	0.30	N
Birth weight	0.41	0.37	N
Mode of delivery*	0.40	0.42	N
Maternal race[a]	<0.01	<0.01	Y
Maternal ethnicity[a]	0.90	<0.01	N
Maternal educational attainment[b]	0.96	<0.01	N
Median household income[b]	0.82	<0.01	N
Maternal marital status	0.32	0.70	N
WIC use[b]	0.18	0.02	Y
Percent cropland within 500m of home[c]	0.24	<0.01	N
Annual PM 2.5 [ug/m^3][c]	0.93	<0.01	N

Data are complete for all variables. P-values generated using χ^2 tests. WIC: Supplemental Nutritional Program for Women, Infants, and Children. *Included in highly adjusted model due to known biological importance. [a]Maternal race and ethnicity are highly correlated. [b]Maternal educational attainment, median household income, and use of WIC all represent socio-economic status and are highly correlated with one another. [c]T-test p-value.

Table S2. χ^2 p-values representing relationships between potential covariates and confounders considered for model inclusion.

Covariate	Gestational age	Weight	Mode of delivery	Maternal race	Maternal ethnicity	Maternal educational attainment	Median household income	Maternal marital status	WIC use	Percent cropland within 500m of home (quartile)[a]	Annual PM 2.5 (quartile [ug/m³])[a]
Sex	0.84	<0.01	0.08	0.88	0.75	0.44	0.66	0.97	0.48	0.25	0.44
Gestational age	-	<0.01	<0.01	0.01	0.75	0.35	0.35	<0.01	0.27	0.03	0.02
Weight		-	<0.01	<0.01	0.02	0.39	0.01	<0.01	0.07	<0.01	0.53
Mode of delivery			-	0.13	0.58	0.02	0.59	0.05	0.10	0.58	<0.01
Maternal race				-	<0.01	<0.01	<0.01	<0.01	<0.01*	<0.01	<0.01
Maternal ethnicity					-	<0.01	<0.01	<0.01	<0.01	<0.01	0.16
Maternal educational attainment						-	<0.01	<0.01	<0.01	<0.01	0.32

Median household income							-	<0.01	<0.01	<0.01	<0.01
Maternal marital status								-	<0.01	<0.01	0.25
WIC use									-	<0.01	0.09
Percent cropland within 500m of home (quartile)[a]										-	<0.01

Data are complete for all variables. P-values generated using χ^2 tests. WIC: Supplemental Nutritional Program for Women, Infants, and Children. *Despite being statistically significantly associated, statistical checks revealed no collinearity or convergence issues when included together in the parsimonious model.
[a]Assessed as quartiles.

Sensitivity analysis using the highly adjusted model.

Table S3. Table of the risk of childhood acute lymphoblastic leukemia (odds ratios and 95% confidence intervals) by buffer size for the highly adjusted model* assessed with the *aggregate and water-specific metric* for the primary and perinatal exposure windows.

Exposure metric and buffer size	OR (95% CI)
Primary window	
ID^2W	
2 km	1.70 (0.91, 3.23)
5 km	1.18 (0.76, 1.76)
10 km	1.03 (0.75, 1.40)
ID_{ups}	
2 km	1.67 (0.64, 4.35)
5 km	1.23 (0.64, 2.38)
10 km	1.09 (0.64, 1.86)
Perinatal window	
ID^2W	
2 km	2.27 (0.89, 5.79)
5 km	1.36 (0.79, 2.34)
10 km	1.30 (0.90, 1.88)
ID_{ups}	
2 km	2.37 (0.55, 10.10)
5 km	1.24 (0.49, 3.12)
10 km	1.33 (0.69, 2.59)

Odds ratios and 95% confidence intervals calculated using unconditional logistic regression. *Highly adjusted: adjusted for year of birth, maternal race, WIC, delivery route, sex, birth weight, percent cropland. OR: Odds Ratio; 95% CI: Confidence Interval.