


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

Cassandra J. Clark,¹  Nicholas P. Johnson,^{1,2} Mario Soriano Jr.,³ Joshua L. Warren,⁴ Keli M. Sorrentino,² Nina S. Kadan-Lottick,⁵ James E. Saiers,³ Xiaomei Ma,⁶ and Nicole C. Deziel^{1,2}

¹Department of Environmental Health Sciences, Yale School of Public Health, Yale University, New Haven, Connecticut, USA

²Center for Perinatal, Pediatric and Environmental Epidemiology, Yale University Schools of Public Health and Medicine, New Haven, Connecticut, USA

³Yale School of the Environment, Yale University, New Haven, Connecticut, USA

⁴Department of Biostatistics, Yale School of Public Health, Yale University, New Haven, Connecticut, USA

⁵Georgetown Lombardi Comprehensive Cancer Center, Washington, District of Columbia, USA

⁶Department of Chronic Disease Epidemiology, Yale School of Public Health, Yale University, New Haven, Connecticut, USA

BACKGROUND: Unconventional oil and gas development (UOGD) releases chemicals that have been linked to cancer and childhood leukemia. Studies of UOGD exposure and childhood leukemia are extremely limited.

OBJECTIVE: The objective of this study was to evaluate potential associations between residential proximity to UOGD and risk of acute lymphoblastic leukemia (ALL), the most common form of childhood leukemia, in a large regional sample using UOGD-specific metrics, including a novel metric to represent the water pathway.

METHODS: We conducted a registry-based case–control study of 405 children ages 2–7 y diagnosed with ALL in Pennsylvania between 2009–2017, and 2,080 controls matched on birth year. We used logistic regression to estimate odds ratios (ORs) and 95% confidence intervals (CIs) for the association between residential proximity to UOGD (including a new water pathway-specific proximity metric) and ALL in two exposure windows: a primary window (3 months preconception to 1 y prior to diagnosis/reference date) and a perinatal window (preconception to birth).

RESULTS: Children with at least one UOG well within 2 km of their birth residence during the primary window had 1.98 times the odds of developing ALL in comparison with those with no UOG wells [95% confidence interval (CI): 1.06, 3.69]. Children with at least one vs. no UOG wells within 2 km during the perinatal window had 2.80 times the odds of developing ALL (95% CI: 1.11, 7.05). These relationships were slightly attenuated after adjusting for maternal race and socio-economic status [odds ratio (OR) = 1.74 (95% CI: 0.93, 3.27) and OR = 2.35 (95% CI: 0.93, 5.95)], respectively. The ORs produced by models using the water pathway-specific metric were similar in magnitude to the aggregate metric.

DISCUSSION: Our study including a novel UOGD metric found UOGD to be a risk factor for childhood ALL. This work adds to mounting evidence of UOGD's impacts on children's health, providing additional support for limiting UOGD near residences. <https://doi.org/10.1289/EHP11092>

Introduction

Childhood acute lymphoblastic leukemia (ALL) is a hematological malignancy that arises from immature B- and less commonly T-lymphoid immune cells.¹ ALL is the most common type of cancer in children (age 0–14 y), representing nearly 80% of childhood leukemia cases and 20%–30% of all childhood cancer cases.^{1–3} Incidence of ALL typically peaks in children age 2–4 y,^{1,4} indicating that the early life environment is likely etiologically important. Although long-term survival rates exceed 90%,⁵ survivors may face health and wellness difficulties later in life, such as chronic illnesses (e.g., cognitive dysfunction, heart disease),^{6–9} psychological issues (e.g., depression, anxiety),^{9–11} and elevated risk of second primary cancers.⁸ Despite a decrease in the incidence of cancer overall in the United States, the incidence of childhood ALL has continued to increase, underscoring the importance of primary prevention.

The etiology of ALL is likely multifactorial and attributable to both environmental exposures and underlying genetic susceptibility. Current evidence suggests that for most cases, ALL develops due to multiple genetic insults, such as chromosomal translocations or alterations.^{12–14} The development of preleukemic clone cells commonly occurs after an initiating genetic insult from a

chromosomal translocation *in utero*, with an additional genetic insult required for overt ALL to manifest.^{2,4,14,15} Although the genetic and molecular processes behind the disease have been delineated, the upstream etiological agents triggering such biological insults remain poorly understood. Current evidence and the early age of peak ALL incidence suggest that exposure to environmental chemicals—particularly to chemicals that are hematotoxic, damage DNA, or interfere with the immune system—may provide a mechanism for pre- or postnatal insults.^{2,16} To date, ALL has been linked to several environmental and chemical exposures, including ionizing or diagnostic radiation,^{17,18} radon,¹⁹ air pollution,^{20–24} pesticides,^{25–29} polybrominated diphenyl ethers,³⁰ and benzene.^{22,31–35}

Unconventional oil and gas development (UOGD), commonly referred to as hydraulic fracturing or “fracking,” is a complex process with the potential for releases of chemical and radiological contaminants into both water and air.³⁶ UOGD is a rapidly expanding source of energy and petrochemical production in the United States. Hydraulic fracturing, an important step in the UOGD process, involves pressurized injections of millions of gallons of water, chemicals, and proppant (e.g., sand) into underground rock formations to create small fissures, allowing natural gas to flow to the surface.³⁷ In addition to the natural gas, the injected fluids and formation water also rise to the surface as wastewater. A single well has been estimated to produce between 1.7 and 14 million liters of wastewater over the first 5 to 10 y of production, and this varies widely by producing formation.^{38,39} The transport and storage of this wastewater may result in surface spills,^{40–43} and improper management or structural failures of injection wells used for storage can result in migration of chemicals into groundwater or surface water.^{44–46} Average annual spill rates (number of spills/UOG wells drilled) across four states was estimated at 5.6%, with 31.1% of wells ever reporting a spill; many spills occurred in watersheds serving as drinking water sources.⁴²

Address correspondence to Cassandra J. Clark. Email: cassie.clark@yale.edu

Supplemental Material is available online (<https://doi.org/10.1289/EHP11092>).

The authors declare they have nothing to disclose.

Received 11 February 2022; Revised 15 June 2022; Accepted 29 June 2022; Published 17 August 2022.

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 ehpsubmissions@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

Hundreds of chemicals have been reportedly used in UOGD injection water or detected in wastewater, some of which have been associated with leukemia.⁴⁷ Known and suspected carcinogens include heavy metals, radioactive material, volatile organic compounds (e.g., benzene), and polycyclic aromatic hydrocarbons.^{48,49} In addition to water pollution, UOGD has the potential to generate air pollution during well and road construction and through vehicle emissions from the transport of oil, gas, and wastewater.^{50,51} Studies of UOGD-related air emissions have measured several carcinogens, including radioactivity, particulate matter (PM), and volatile organic compounds (e.g., benzene).^{52–55} Furthermore, elevated levels of indoor radon were measured in homes near UOGD activity.^{56,57} Additionally, the process of extracting natural gas also brings technologically enhanced naturally occurring radioactive compounds to the surface with ancient brine formation water, and drill cuttings and sludge from equipment may also contain radioactivity.^{58,59} The potential for children living near UOGD to be exposed to chemical carcinogens and radiological contaminants is a major public health concern.

Research on the potential association between exposure to UOGD and risk of childhood cancer is urgently needed. To our knowledge, there have been only two published studies of this relationship to date. The first was an ecological study conducted in the state of Pennsylvania,⁶⁰ which compared standardized incidence ratios of childhood cancer before and after drilling and observed no difference; this analysis did not account for a latency period or adjust for confounders.⁶¹ The second, a registry-based case–case study in Colorado, found that children and young adults with ALL (ages 0–24 y; *n* = 87 cases) were four times more likely to live in areas of greater oil and gas activity (conventional and unconventional combined) than controls, which were children with

nonhematological cancers, based on models adjusted for multiple confounders.⁶² The case–case methodology may have attenuated the true association if UOGD was a shared risk factor. The paucity of data on the association between UOGD and childhood cancer outcomes has fueled public concerns about possible cancer clusters in heavily drilled regions and calls for more research and government action.⁶³

To advance understanding of the relationship between UOGD exposure and ALL risk and inform public policy, we conducted a registry- and population-based case-control study. This work builds on prior studies by incorporating a larger sample size, the use of cancer-free controls identified from birth records, and the use of UOGD-specific metrics, including a novel metric developed for capturing exposures through the water pathway.^{64,65}

Methods

Study Setting, Population, and Design

We conducted a population-based case–control study in the commonwealth of Pennsylvania because it is home to intense oil and gas activity. More than 10,000 UOG wells were drilled in Pennsylvania between 2002 and 2017, with the pace of drilling increasing sharply from 2007 to 2011.⁶⁶ In addition, more than 1,000 spills, 5,000 violations, and 4,000 resident complaints related to oil and gas were documented between 2005 and 2014 in Pennsylvania.^{42,67} Further, up to one-third of domestic groundwater wells in Pennsylvania are located within 2 km of a hydraulically fractured well.⁶⁸

Cases included all children diagnosed with ALL between the ages of 2–7 y in Pennsylvania from 2009 to 2017 (Figure 1). We chose this age range to cover the peak age of ALL incidence in

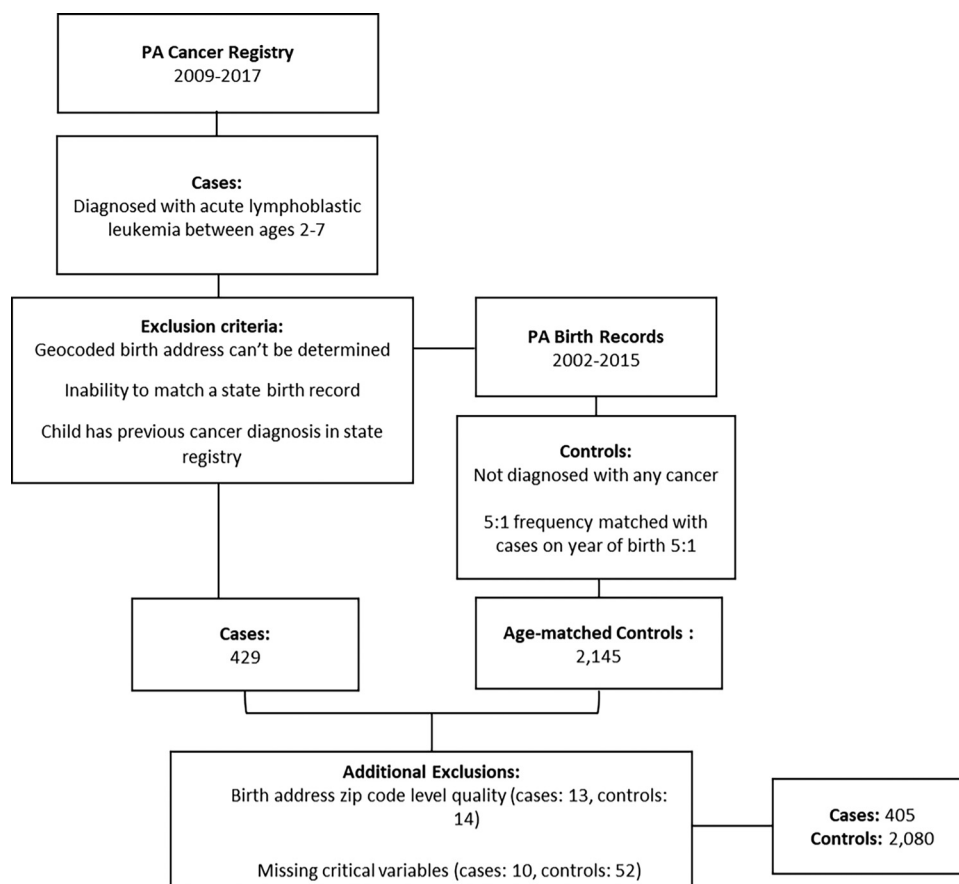


Figure 1. Data sources and selection process for Pennsylvania cases and controls (2009–2017).

the United States⁶⁹ and exclude cases of the etiologically distinct infant leukemia (diagnosis between the ages of 0–1 y).^{70,71} We selected the years of diagnosis to ensure there was opportunity for exposure after drilling commenced in the state and a latency period of at least 1 y to account for the development of disease.⁷² ALL cases ($n = 429$) were identified from the Pennsylvania state cancer registry by Pennsylvania Department of Health staff using ICD-O-3 sites C420, C421, C424 and Histology codes 9811–9818, 9826, and 9835–9837. Cases were then linked to their birth records available from the Pennsylvania Vital Records maintained by the Bureau of Health Statistics and Registries. Cases were excluded if *a*) the state could not match a birth record in Pennsylvania, *b*) the child had a previous diagnosis of cancer in the state cancer registry, and *c*) a birth address could not be obtained/geocoded beyond ZIP code level.

For each case, five control children were randomly selected by Pennsylvania Department of Health staff from live births in the Pennsylvania birth records with frequency-matching on birth year ($n = 2,145$; Figure 1). Reasons for excluding controls included: *a*) birth address could not be obtained or geocoded to street level, *b*) the child had a previous diagnosis of cancer in the state cancer registry, and *c*) the child was a sibling of a case or another control. After obtaining the data set, we performed additional geocoding [using SAS (version 9.4; SAS Institute Inc.) and checked geocode quality for both case and control children, excluding those whose birth address was not street-level quality or better ($n = 14$ cases; $n = 13$ controls). Because the missingness rate for several key covariates was very low, we elected to conduct a complete case analysis by excluding children from the study population missing the following covariates (established or suspected risk factors)^{16,73–75}: maternal participation in the U.S. Department of Agriculture’s Special Supplemental Nutrition Program for Women, Infants, and Children (WIC, an individual-level representation of socioeconomic status), birth weight, and mode of delivery (Figure 1; $n = 10$ cases; $n = 52$ controls). We included 405 cases and 2,080 controls in our final analyses. The study protocol was approved by the institutional review board of Yale University (HIC #2000021809) and by the Pennsylvania Department of Health.

Exposure Assessment

We obtained and merged permit and production report data sets from the Pennsylvania Department of Environmental Protection’s Office of Oil and Gas Management⁶⁶ to construct a data set of location, permit, and production data for UOG wells that were active (i.e., drilled or producing, as confirmed by having a reported spud date or a submitted production report) in Pennsylvania during the period 2001–2015. The data were then cleaned, and their quality were checked. For example, missing data on spud date, well type, and producing formation in the permit data sets were cross-referenced with and supplemented by the production data sets. Duplicate entries were addressed by preferentially retaining the most recent entry. Wells with a missing spud date were assigned a spud date equal to the first date of the earliest production report minus the median number of days between spud and first production in the data set. The final database included 9,578 active coalbed methane, gas, oil, and combined oil and gas wells in unconventional formations.

Maternal residential address at birth was obtained and geocoded from birth records for both cases and controls, and address at diagnosis was obtained from cancer registries for cases. Birth address was used to assign exposures using inverse distance-squared weighted (ID²W) well counts (represented by $\sum_{i=1}^n \frac{1}{d_i^2}$ for all UOG wells within a buffer zone, where d is distance between

the i^{th} UOG well and a residence), referred to as the “aggregate metric.” We calculated this metric with buffer sizes of 2, 5, and 10 km. We selected two etiologically important exposure windows: *a*) 3 months prior to conception to 1 y prior to diagnosis, called the “primary window,” and *b*) 3 months prior to conception to birth, called the “perinatal window.” For the primary window, age-matched controls were assigned a reference date corresponding to the diagnosis date of a case. For the perinatal window, exposures were assigned using the respective birth dates of the cases and controls.

To capture water as a route of exposure to UOGD, we also calculated a flow-direction metric based on land-surface topography, inverse distance metric ID_{ups}, referred to as the “water pathway-specific metric.” ID_{ups} is based on the widely accepted conceptual model that groundwater flow in regions of hill-and-valley topography occurs in the downhill direction, parallel to the topographic gradient.⁷⁶ ID_{ups} is represented by the equation $\frac{1}{u}$, where (u) is distance to the nearest upgradient UOG well, determined with the D-infinity algorithm in TauDEM. This metric and its underlying programming code was introduced by Soriano et al. and was subsequently applied in a study of UOGD-related drinking water exposure.⁶⁵ This exposure metric assumes that UOG wells that are located upgradient of a residence contribute more to exposure than downgradient wells, presuming that consumption or contact with groundwater from domestic wells is a major exposure source. The metric was calculated using buffer sizes of 2, 5, and 10 km around the maternal residence. Our selection of buffer sizes was informed by the hydrological (2 km) and epidemiological (5 and 10 km) literature,^{64,65,76,77–81} and facilitates comparison between the aggregate metric and the water pathway-specific metric and comparisons with previous epidemiologic studies. We conducted a subanalysis using this metric as the main UOGD exposure assessment variable.

Residential Mobility

Residential mobility among pregnant women or in early childhood could introduce exposure misclassification.^{82–85} We used three analyses to address the potential exposure misclassification introduced by residential mobility among pregnant mothers. First, we compared all case addresses at birth and diagnosis and assessed the distance moved as well as variables associated with mobility (e.g., socioeconomic status). Second, we examined the difference in cases’ exposure classification at the birth and diagnosis addresses. Third, our selection of the perinatal exposure window, which restricts the window of exposure to 3 months prior to conception to birth, addresses mobility. The exposure estimate based on birth address is likely to be most accurate during this shorter time window (i.e., less opportunity to move residences), and pregnancy is an important etiological window for childhood leukemia.^{14,15,29} We used the findings from these three analyses to provide context and aid interpretation of our results.

Covariates and Confounders

To account for potential confounding, we considered adjustment for both individual-level and area-level factors. We generated a list of *a priori* potential confounders informed by the literature that were available from birth records or publicly available data sources including sex, mode of delivery, birth weight, race, ethnicity, maternal education, air pollution exposure, and pesticide exposure.^{2,20,21,25,33,73–75,86–88}

We estimated exposure to maternal and childhood residential air pollution using the U.S. Environmental Protection Agency (U.S. EPA) Bayesian space–time downscaler models, which provide daily estimates of average fine PM with an aerodynamic diameter

≤2.5 μm (PM_{2.5}) at census tract centroids.⁸⁹ We took the mean of daily average PM_{2.5} measurements from 3 months prior to conception to 1 y prior to diagnosis to produce one representative PM_{2.5} measurement for each individual. To represent maternal and childhood residential exposure to agricultural pesticides, we retrieved raster data of cropland for Pennsylvania from the U.S. Department of Agriculture National Agricultural Statistics Service CropScape.⁹⁰ Individuals were matched to the cropland map from their birth year, except for 2003 and 2004, which used a 2002 map, and 2005–2008, which used a 2008 map, due to data availability. We calculated the percent of land designated as cropland within buffers of 500 m and 1,000 m around each home (modeled after Reynolds et al.⁹¹ and referred to this as “percent cropland”).

We obtained information on community-level demographic and socioeconomic characteristics from the U.S. 2000 and 2010 Decennial Census [e.g., median household income, educational attainment, percentage of households living in poverty, housing

occupancy, housing type (e.g., rented vs. owned)] for all Pennsylvania census tracts.⁹² We also linked individuals to the Centers for Disease Control and Prevention/Agency for Toxic Substances and Disease Registry Social Vulnerability Index (SVI), a composite metric representing 15 different social conditions, including socioeconomic status, demographics, and access to transportation, among other factors.⁹³

Statistical Analyses

All statistical analyses were conducted in SAS (version 9.4; SAS Institute Inc.), and all tests were two-sided with an alpha level of 0.05. We used unconditional logistic regression to estimate odds ratios (ORs) and 95% confidence intervals (CIs) for the association between UOGD exposure and ALL risk, adjusting for year of birth (i.e., the matching variable). We constructed separate models for each metric, for buffer size, and for both the primary

Table 1. Distribution of Pennsylvania study population characteristics (2009–2017).

Variable	Cases (n = 405)	Controls (n = 2080)	χ ² p-value
	n (%)	n (%)	
Sex	—	—	0.57
Male	222 (55)	1,108 (53)	—
Female	183 (45)	972 (47)	—
Gestational age (wk)	—	—	0.76
<32 wk (Very preterm)	5 (1)	40 (2)	—
32 to <37 (Preterm)	35 (9)	162 (8)	—
37 to <39 (Early term)	78 (19)	436 (21)	—
39–41 (Term)	258 (64)	1,275 (61)	—
42+ (Postterm)	28 (7)	155 (7)	—
Out of limit, missing, no physician estimate	1 (1)	12 (1)	—
Birth weight	—	—	0.41
Low birth weight (<2,499 g)	27 (7)	172 (8)	—
Normal birth weight (2,500–3,999 g)	333 (82)	1,707 (82)	—
High birth weight (>4,000 g)	45 (11)	201 (10)	—
Delivery route	—	—	0.40
Vaginal	281 (69)	1,399 (67)	—
Cesarean	124 (31)	681 (33)	—
Mother’s race	—	—	<0.0001
White	327 (81)	1,520 (73)	—
Black	29 (7)	333 (16)	—
Other	42 (10)	179 (9)	—
Not reported	7 (2)	48 (2)	—
Mother’s ethnicity	—	—	0.90
Not Hispanic	370 (91)	1,888 (91)	—
Hispanic	31 (8)	173 (8)	—
Unknown	4 (1)	19 (1)	—
Mother’s educational attainment	—	—	0.96
High school or less	54 (13)	266 (13)	—
Some college	221 (55)	1,129 (54)	—
Bachelor’s	84 (21)	430 (21)	—
>16 y	46 (11)	255 (12)	—
Mother uses WIC	—	—	0.18
Yes	160 (40)	749 (36)	—
No	245 (60)	1,331 (64)	—
Median household income (\$USD)	—	—	0.88
<\$26,500	96 (24)	517 (25)	—
\$26,500–\$53,000	191 (47)	971 (47)	—
>\$53,000	118 (29)	492 (28)	—
	Mean (SD)	Mean (SD)	—
Percent cropland (500 m ^a)	—	—	0.24 ^b
	13.8 (20.9)	12.5 (20.4)	—
	—	—	0.71 ^b
CDC SVI percentile	54.0 (27.9)	53.4 (29.7)	—
Annual PM _{2.5} (μg/m ³)	—	—	—
Primary window	11.7 (1.7)	11.7 (1.7)	0.93 ^b
Perinatal window	12.4 (2.1)	12.4 (2.2)	0.91 ^b

Note: Data are complete for all variables. p-Values generated using χ² tests. —, no data; CDC, U.S. Centers for Disease Control and Prevention; IQR, interquartile range; SVI, CDC/Agency for Toxic Substances and Disease Registry Social Vulnerability Index; USD, United States dollars; WIC, Supplemental Nutritional Program for Women, Infants, and Children.

^aUsed as a proxy for pesticide exposure, accounting for likely extent of pesticide drift; calculated for year of birth only.

^bt-test p-value.

and perinatal exposure windows. We constructed two model types: minimally adjusted (i.e., only adjusting for year of birth via matching) and parsimonious (i.e., only covariates that changed the OR by 10% or more) (see Supplemental Material, “Intermediate analyses of association and correlation to identify covariates and confounders for model building”). If two covariates were highly correlated (Spearman $\rho > 0.80$ or $\chi^2 p < 0.05$) and led to model convergence problems, one was selected for use based on public health relevance (e.g., though both representing socioeconomic status, an individual-level measure of socioeconomic status such as maternal use of food stamps may be more relevant to a child’s health outcome than their census tract-level median household income) and distribution in the population (e.g., heterogeneity of exposure). We considered several individual- and community-level variables that are proxy measures of socioeconomic status, including maternal education, maternal participation in WIC, census tract-level median household income, and census tract-level SVI. The parsimonious models included maternal race and maternal participation in WIC. As a sensitivity analysis, we constructed a third highly adjusted model that included those covariates that were either associated with the exposure or outcome based on χ^2 and Fisher’s exact tests at a less stringent $p < 0.20$ (Supplemental Material, “Intermediate analyses of association and correlation to identify covariates and confounders for model building”) or had known etiological or biological importance according to the literature (infant sex, mode of delivery).

Results

Demographics

Cases and controls were similar with respect to sex, gestational age, birth weight, mode of delivery, educational attainment of the mother, census tract-level median household income, and SVI (Table 1). Mothers were predominantly non-Hispanic (91% of both cases and controls) and White, but there was a higher percentage of White mothers among cases (81% of cases and 73% of controls). The case group had a significantly smaller percentage of Black

Table 2. Exposure prevalence in 405 childhood acute lymphoblastic leukemia cases and 2,080 age-matched controls across exposure windows, metrics, and buffer sizes.

Exposure metric and buffer size	Primary window		Perinatal window	
	Cases (n = 405)	Controls (n = 2,080)	Cases (n = 405)	Controls (n = 2,080)
	n (%)	n (%)	n (%)	n (%)
ID ² W 2 km				
Exposed	14 (3)	37 (2)	7 (2)	13 (1)
Unexposed	391 (97)	2,043 (98)	398 (98)	2,067 (99)
ID ² W 5 km				
Exposed	31 (8)	122 (6)	18 (4)	61 (3)
Unexposed	374 (92)	1,958 (94)	387 (96)	2,019 (97)
ID ² W 10 km				
Exposed	59 (15)	270 (13)	41 (10)	153 (7)
Unexposed	346 (85)	1,810 (87)	364 (89)	1,927 (83)
ID _{ups} 2 km				
Exposed	6 (2)	16 (1)	3 (1)	5 (1)
Unexposed	399 (98)	2,064 (99)	402 (99)	2,075 (99)
ID _{ups} 5 km				
Exposed	12 (3)	43 (2)	6 (1)	21 (1)
Unexposed	393 (97)	2,037 (98)	399 (99)	2,059 (99)
ID _{ups} 10 km				
Exposed	18 (5)	74 (4)	12 (3)	39 (2)
Unexposed	346 (95)	1,810 (96)	393 (97)	2,041 (98)

Note: Exposure for each buffer size and metric was dichotomized due to low exposure prevalence. ID²W, inverse distance-squared weighted well count; ID_{ups}, inverse distance to the nearest upgradient UOG well; UOG, unconventional oil and gas.

mothers (7% in comparison with 16% of controls). A slightly greater frequency of mothers of cases reported participating in WIC (40% of cases and 36% of controls). Case children had a greater percentage of cropland within 500 m of their birth address on average than control children (13.8% in comparison with 12.5%). Average annual PM_{2.5} levels were not significantly different between cases and controls (11.7 $\mu\text{g}/\text{m}^3$ for both groups).

UOGD Exposure within the Study Population

A total of 85%–98% of the study population was unexposed to UOGD; the prevalence of unexposed varied based on exposure metric buffer sizes (Table 2). Due to the low prevalence and limited variability in UOGD exposure, we dichotomized our exposure assessment metrics, because there was insufficient spread to apply them with more than two categories or use them continuously. The ID²W metric, when dichotomized, effectively represents whether the participant had at least one UOG well within the buffer zone, whereas the ID_{ups} metric represents whether the participant had at least one UOG well within the buffer zone that was located upgradient within their watershed.

Residential Mobility

A total of 58% of cases moved residences between birth and diagnosis. The mean distance moved was 9.02 km (median: 0.49 km, interquartile range: 0–4.88 km, range: 0–374 km). Though the proportion of cases who moved (and for some, the distance moved) was substantial, <2% of individuals changed exposure designation (either exposed to unexposed or vice versa) using any metric after the move.

Association between ALL and Exposure to UOGD

Aggregate metric (ID²W). Using the aggregate UOG exposure metric and the primary exposure window, ORs were elevated for individuals living within 2, 5, and 10 km of UOGD (Figure 2). In models adjusting only for year of birth, the odds of developing ALL were 1.98 times higher in children with at least one UOG well within 2 km of their birth residence, in comparison with those with no UOG wells (95% CI: 1.06, 3.69). The magnitude of the minimally adjusted OR decreased monotonically but remained elevated as the buffer size of the exposure metrics increased to 5 km (OR = 1.33; 95% CI: 0.88, 2.00) and 10 km (OR = 1.14; 95% CI: 0.84, 1.55). After adjusting for maternal race and WIC participation in our parsimonious models, the odds of ALL were 1.74 times higher for individuals living within 2 km of UOGD (95% CI: 0.93, 3.27), with some attenuation of the odds ratio at buffer sizes of 5 km (OR = 1.18; 95% CI: 0.78, 1.78) and 10 km (OR = 1.03; 95% CI: 0.75, 1.40). Our sensitivity analysis, which included adjustment for the additional covariates of sex, delivery route, birth weight, and percentage cropland, did not appreciably change the estimates in comparison with the parsimonious model (Supplemental Material, “Sensitivity analysis using the highly adjusted model”).

For the aggregate metric and the perinatal window, estimates were larger in magnitude by 20%–40% than the estimate for the corresponding buffer size using the primary window (Figure 2). Children living within 2 km of UOGD had 2.80 times the odds of developing ALL (95% CI: 1.11, 7.05) in models adjusting only for year of birth. The minimally adjusted odds of ALL were also elevated for children with UOGD within 5 km (OR = 1.54; 95% CI: 0.90, 2.63) and 10 km (OR = 1.42; 95% CI: 0.99, 2.04). In parsimonious models, children with UOGD within 2 km had 2.35 times the odds of having ALL (95% CI: 0.93, 5.95). In sensitivity analyses, the highly adjusted model results were consistent with the parsimonious models at all buffer sizes (Supplemental Material, “Sensitivity analysis using the highly adjusted model”).

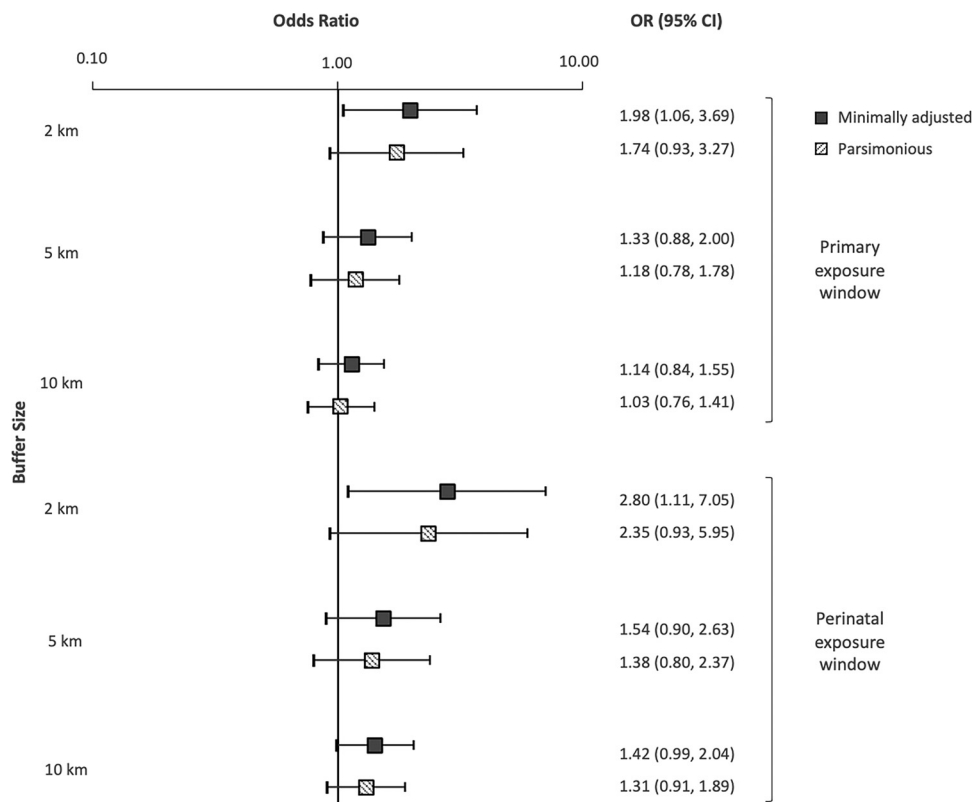


Figure 2. Plots of the risk of childhood acute lymphoblastic leukemia (ORs and 95% CIs) by buffer size, assessed with the aggregate metric for the primary and perinatal exposure windows. The aggregate metric refers to ID²W well counts. ORs and 95% CIs calculated using unconditional logistic regression. Minimally adjusted: adjusted for year of birth only; Parsimonious: adjusted for year of birth, maternal race, and WIC. Note: CI, confidence interval; ID²W inverse distance-squared weighted; OR, odds ratio; WIC, Supplemental Nutritional Program for Women, Infants, and Children.

Water pathway-specific metric. Use of the water pathway-specific exposure metric in the regression models produced results that were similar to those for the aggregate metric for the primary exposure window (Figure 3). Children who had at least one upgradient UOG well within 2 km had 1.94 times the odds of developing ALL (95% CI: 0.75, 4.99) in comparison with unexposed children in models adjusting only for year of birth, though the CI was wide. The association was slightly attenuated by adjusting for maternal race and WIC participation (OR = 1.70; 95% CI: 0.66, 4.41), and the most adjusted model results were consistent with the parsimonious model. Children with at least one upgradient UOG well within 5 km had 1.45 times the odds of developing ALL (95% CI: 0.76, 2.77). Finally, children with at least one upgradient UOG well within 10 km in their watershed had 1.26 times higher odds of developing ALL than unexposed children (95% CI: 0.75, 2.14). Adjusting for maternal race and WIC participation attenuated this association (OR = 1.10; 95% CI: 0.64, 1.87). The estimates produced by the sensitivity analyses were not appreciably different from those produced by the parsimonious model (Supplemental Material, “Sensitivity analysis using the highly adjusted model”).

The ORs for the water pathway-specific metric restricted to the perinatal window were also similar to those produced by the aggregate metric (Figure 2). In models adjusting only for year of birth, children with UOG activity within 2 km falling within their upgradient watershed had 3.10 times the odds of developing ALL (95% CI: 0.74, 13.01). In the parsimonious model, the odds of developing ALL for those children were 2.45 (95% CI: 0.58, 10.37). Children with an upgradient UOG well within 5 and 10 km had 1.48 and 1.60 times higher odds, respectively, of developing ALL than control children (95% CI: 0.59, 3.68 and 0.83,

3.08, respectively). The odds remained elevated at 5 and 10 km in the parsimonious model. Sensitivity analyses adjusting for additional covariates including sex, delivery route, birth weight, and percentage cropland did not significantly change the estimates in comparison with the parsimonious model (Supplemental Material, “Sensitivity analysis using the highly adjusted model”).

Discussion

In this population-based case-control study of UOGD including 405 children with ALL and 2,080 age-matched controls in Pennsylvania, we found that children living in proximity to UOGD had up to 2–3 times the odds of developing ALL. Although ORs were statistically significant in models only accounting for year of birth, elevated ORs persisted after additionally adjusting for race, socioeconomic status, and competing environmental exposures. However, low exposure prevalence limited our statistical power, and confidence intervals at the 2 km buffer size and for the water pathway-specific metric in particular were wide. Nonetheless, our results indicate that exposure to UOGD may be an important risk factor for ALL, particularly for children exposed *in utero*. To our knowledge, this is the first case-control study of childhood ALL that examined UOGD exposure exclusively, the largest study of unconventional oil and gas and hematological malignancies in children, and the first study to apply a water pathway-specific metric of UOGD exposure in a health context.

Our results complement those reported by the McKenzie et al. study in Colorado, which reported significantly elevated odds of ALL for children and young adults ages 5–24 y and nonsignificantly elevated or mixed odds for children ages 0–4 y.⁶² In the Colorado study, the strongest odds were observed for children and young adults ages 5–24, who were 3–4 times as likely to live

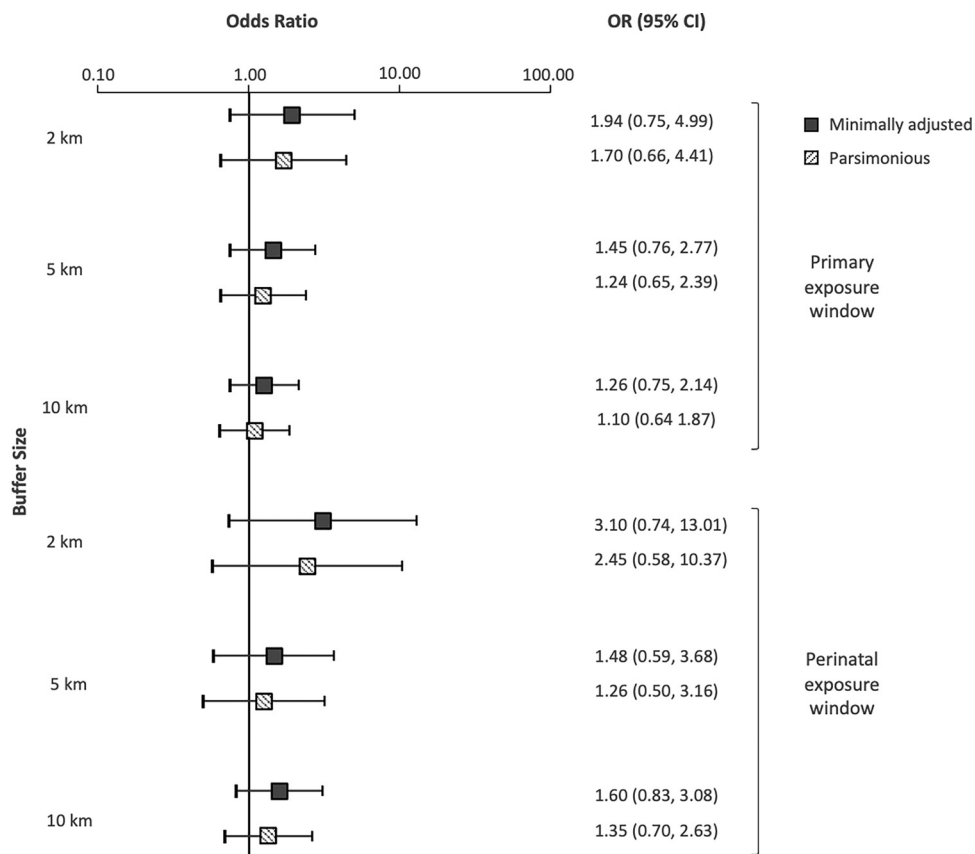


Figure 3. Plots of the risk of childhood acute lymphoblastic leukemia (ORs and 95% CIs) by buffer size, assessed with the water pathway-specific metric for the primary and perinatal exposure windows. ORs and 95% CIs calculated using unconditional logistic regression. Minimally adjusted: adjusted for year of birth only; Parsimonious: adjusted for year of birth, maternal race, and WIC. Note: CI, confidence interval; OR, odds ratio; WIC, Supplemental Nutritional Program for Women, Infants, and Children.

near UOG as control children with nonhematological cancers. Our ORs fell within a similar range. However, their study only had 39 cases in the 0–4 y age range, which may have hindered their ability to draw inferences for that group. Our results, which focus on children ages 2–7 y, provide more information on this younger age group.

Our results also suggest that preconception to birth is an important etiological window for exposure to UOG and the development of ALL. This finding is consistent with research on other environmental exposures, such as pesticides,^{25,26,94} bolstering the evidence for the importance of this sensitive window. ORs calculated using the perinatal window were 20%–40% larger than the estimates for the same buffer size using the primary window, though there were fewer exposed individuals and more uncertainty overall. The perinatal period is a critical window for the genetic mutations that precede the development of ALL.^{14,15} It is generally hypothesized that the etiology of childhood ALL is multifactorial due to two distinct genetic “hits.”⁹⁵ The development of preleukemic clone cells commonly occurs after a genetic insult that results in fusion gene formation or hyperdiploidy *in utero*.^{2,95,96} Then a second, possibly postnatal, insult is required for overt ALL to develop.^{2,4,14,15} Given the similar results observed across both exposure windows, our findings suggest that UOG-related environmental exposures may contribute to both prenatal and postnatal insults leading to the development of ALL.

We applied a new metric for evaluating drinking water exposures from UOGD and identified suggestive relationships between ID_{ups} and ALL. This metric and our selection of buffer sizes were informed by the hydrological and epidemiological

literature.^{64,65,76,77–81} The estimates generated using the water pathway-specific metric ID_{ups} were similar or greater in magnitude in comparison with estimates using the traditional ID²W metric, although the uncertainty associated with these estimates was higher. This finding could indicate that water is an important route of exposure to leukemogenic compounds for the development of ALL. Our metrics do not identify specific etiological agents underlying the observed associations. Seventeen compounds used or produced by UOGD have been previously associated with leukemia.⁴⁷ One candidate agent is benzene. Maternal occupational and ambient exposure to benzene, which is known to be used or produced by UOGD, in the air^{22,32} or in the form of solvents, paints, and petroleum during pregnancy have been associated with elevated odds of ALL.³⁵ Benzene has been detected in multiple groundwater studies in this region focused on UOGD^{41,65,96–99} and in biological samples from communities near oil and gas development.¹⁰⁰ However, it is also possible that these results arose because the water pathway-specific metric produced exposure estimates similar to those of the aggregate metric, particularly when dichotomized. A previous analysis by our group showed that the continuous forms of these metrics tended to be moderately positively correlated with one another (Spearman $\rho = 0.62$ for ID_{ups} and ID²W at 2 km).⁶⁵ It may be that simple proximity to UOGD, which could encompass and/or represent multiple routes of exposure, is the driving factor behind the associations for both metrics. At this time, the dominant stressor is not well understood.³⁶ Nonetheless, epidemiological studies should try to pinpoint specific exposure pathways underlying associations. Several recent studies of UOGD have explored metrics representing specific

(rather than aggregate) routes of exposure, such as flaring, earthquakes, air pollution, and radioactivity.^{50,55,101–103}

Epidemiological studies of UOGD exposure have generally relied on spatial surrogates of exposure, such as ID²W well counts. Previous epidemiological studies using spatial metrics have mainly used a 10 km buffer size or larger.^{104,105} However, when considering environmental exposures like water pollution, realistic transport distances should be considered.¹⁰⁶ A study in northeast Pennsylvania measuring the vulnerability of groundwater wells to contamination by UOGD indicates that the extent of a domestic groundwater well's capture zone (the area around the well from which the water is pulled) is generally less than 2 km.⁷⁷ Further, Llewellyn et al. suggested that a contaminant plume migrated 1 to 3 km in groundwater from a well pad to domestic wells,⁷⁹ and the results of Osborn et al. and Jackson et al. suggest elevated methane levels (i.e., enhanced gas phase transport) within 1 km of UOG well pads.^{107,108} Beyond water, an analysis of UOG-related air pollutants found that individuals whose closest UOG well was <0.5 mi (0.80 km) were at greater risk of health effects from exposure to air pollutants than those further than 0.5 mi from a well.⁵³ The extent of transport of UOG-related air pollutants would be expected to vary by pollutant and local meteorology. Because emitted pollutants attenuate at different functions of distance, there may not be a universal buffer size that optimally captures all hazards. It is possible that applying buffer sizes of 10 km or more could introduce exposure misclassification, dilute the pool of meaningfully exposed individuals, and thus attenuate the magnitude of the observed effect. In this analysis, we observed the largest effect sizes using a buffer size of 2 km, though the number of exposed individuals in these groups was low. The magnitude of the effects at the 5 and 10 km buffer sizes were comparably moderate and was likely particularly apparent in our study because the metrics were used in a binary fashion. Spatial metrics are more typically categorized (e.g., quartiles), and our restricted exposure distribution precluded use of this method. Nonetheless, this attenuation of the observed effect based on the buffer size considered may provide support for using smaller, more selective buffer sizes in epidemiological analyses despite effects to sample size.

This work adds to a growing body of literature on UOGD exposure and women's and children's health used to inform policy, such as setback distances (the required minimum distance between a private residence or other sensitive location and a UOG well).^{109,110} Current setback distances in the United States are the subject of much debate,^{111,112} with some calling for setback distances to be lengthened to more than 305 m (1,000 ft)^{113,114} and as far as 1,000 m (3,281 ft).¹¹⁵ The current setback distance in Pennsylvania is 152 m (500 ft), extended from 61 m (200 ft) in 2012.¹¹⁶ We observed elevated odds of cancer associated with UOG activity within 2 km, which exceeds any existing setback distance. Further, although effect sizes diminished with increasing buffer size, the odds of ALL were still elevated at 5 and 10 km buffer sizes. Our results in the context of the broader environmental and epidemiological literature suggest that existing setback distances are insufficiently protective of public health, particularly for vulnerable populations like children, and should be revisited and informed by more recent data.

Our study has several notable strengths. It is the largest study to date investigating UOGD with ALL or any childhood cancer, the first case-control study to focus exclusively on UOGD exposure, and the first to apply a water pathway-specific UOGD exposure metric. We controlled for multiple known risk factors and examined the impact of several competing environmental exposures. We assessed UOGD exposure at multiple buffer sizes informed by the epidemiological and environmental literature. Selection bias, a

typical concern in case-control studies, is unlikely to have affected our study because we selected cases from population-based cancer registries and controls from statewide birth records without the need to contact any subjects and seek consent for participation. Because we had access to addresses at two time points for the cases, we were able to examine the potential impact of residential mobility on exposure classification. We determined that only a very small percentage of cases (<2%) had different exposure assignments across birth address and diagnosis address, indicating limited potential for exposure misclassification. This finding is consistent with that of other studies of spatially defined environmental exposures, which also have not found residential mobility to be a major source of error.^{82,83}

Our study had several limitations. First, we were constrained by individual-level information available in the birth records, which limited our ability to investigate potential confounders such as parental occupation. Though we designed a statewide study, UOGD is confined to the extent of the shale and drilling is also not performed in urbanized metropolitan centers. Therefore, most of our study population was unexposed. However, we would expect this to have attenuated any observed relationships, because population density and incidence for cancer with nonmodifiable risk factors tends to be higher in urban areas,¹¹⁷ and urban dwelling individuals may be more likely to experience known risk factors for ALL, such as air pollution.^{118,119} Although the ORs were not statistically significant after adjusting for race, socioeconomic status, and other environmental exposures in comparison with ORs from models accounting for year of birth alone, the odds remained consistently elevated across different time periods and metrics. ALL is a rare disease, and as such, the lack of statistical significance could be due more to the rarity of the disease limiting our precision than to lack of biological or public health significance. Low exposure prevalence (between 1% and 5%) when using the water pathway-specific metric (particularly at the smaller buffer sizes) may have reduced model stability and reduced the overall precision of risk estimates. This metric may reveal more differences in larger study populations, or in studies of more common health end points. Moreover, the metric is most relevant for people using private groundwater wells. Although a significant proportion of our suburban and urban population may be served by public water sources, up to 50% of residents in the more heavily drilled rural counties may be served by groundwater wells.^{120,121} There is an opportunity to further examine drinking water sources in this population to improve the accuracy of exposure assessment.

Our study suggests that children living near UOGD have increased odds of developing ALL as assessed by multiple metrics, including a novel metric representing drinking water exposure. The magnitude of the association was greatest among those children living within 2 km of UOGD and exposed during the perinatal period. This research adds to a growing body of work documenting adverse health effects associated with UOGD, particularly among children,^{36,109,110} and provides additional support for more stringent setback policies and other public health measures to reduce exposures to UOGD.

Acknowledgments

This research was supported in part by National Priority Research Project under Assistance Agreement No. CR839249 awarded by the U.S. EPA to Yale University. The publication has not been formally reviewed by the U.S. EPA. The views expressed in this document are solely those of the authors and do not necessarily reflect those of the U.S. EPA. In addition, the U.S. EPA does not endorse any products or commercial services mentioned in this publication. C.J.C. was also supported by the National Institute of Environmental Health Sciences under the

National Institutes of Health (NIH; F31ES031441), the Yale Cancer Center (T32CA250803), and the Yale Institute for Biospheric Studies. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

- Pui C-H. 2011. Acute Lymphoblastic Leukemia. In *Encyclopedia of Cancer*. Schwab M, ed. Berlin, Germany: Springer Berlin Heidelberg, 23–26.
- Eden T. 2010. Aetiology of childhood leukaemia. *Cancer Treat Rev* 36(4):286–297, PMID: 20223594, <https://doi.org/10.1016/j.ctrv.2010.02.004>.
- Siegel DA, Henley SJ, Li J, Pollack LA, Van Dyne EA, White A. 2017. Rates and trends of pediatric acute lymphoblastic leukemia — United States, 2001–2014. *MMWR Morb Mortal Wkly Rep* 66(36):950–954, PMID: 28910269, <https://doi.org/10.15585/mmwr.mm6636a3>.
- Rossig C, Juergens H. 2008. Aetiology of childhood acute leukaemias: current status of knowledge. *Radiat Prot Dosimetry* 132(2):114–118, PMID: 18922818, <https://doi.org/10.1093/rpd/ncn269>.
- Phelan R, Eissa H, Becktell K, Bhatt N, Kudek M, Nuechterlein B, et al. 2019. Upfront therapies and downstream effects: navigating late effects in childhood cancer survivors in the current era. *Curr Oncol Rep* 21(12):104, PMID: 31768799, <https://doi.org/10.1007/s11912-019-0861-8>.
- Bhakta N, Liu Q, Ness KK, Baassiri M, Eissa H, Yeo F, et al. 2017. The cumulative burden of surviving childhood cancer: an initial report from the St Jude Lifetime Cohort Study (SJLIFE). *Lancet* 390(10112):2569–2582, PMID: 28890157, [https://doi.org/10.1016/S0140-6736\(17\)31610-0](https://doi.org/10.1016/S0140-6736(17)31610-0).
- Oeffinger KC, Mertens AC, Sklar CA, Kawashima T, Hudson MM, Meadows AT, et al. 2006. Chronic health conditions in adult survivors of childhood cancer. *N Engl J Med* 355(15):1572–1582, PMID: 17035650, <https://doi.org/10.1056/NEJMs060185>.
- Mertens AC, Liu Q, Neglia JP, Wasilewski K, Leisenring W, Armstrong GT, et al. 2008. Cause-specific late mortality among 5-year survivors of childhood cancer: the Childhood Cancer Survivor Study. *J Natl Cancer Inst* 100(19):1368–1379, PMID: 18812549, <https://doi.org/10.1093/jnci/djn310>.
- Hudson MM, Mertens AC, Yasui Y, Hobbie W, Chen H, Gurney JG, et al. 2003. Health status of adult long-term survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *JAMA* 290(12):1583–1592, PMID: 14506117, <https://doi.org/10.1001/jama.290.12.1583>.
- Baytan B, Aşut Ç, Çirpan Kantarcıoğlu A, Sezgin Evim M, Güneş AM. 2016. Health-related quality of life, depression, anxiety, and self-image in acute lymphocytic leukemia survivors. *Turk J Haematol* 33(4):326–330, PMID: 27094799, <https://doi.org/10.4274/tjh.2015.0356>.
- Kunin-Batson A, Kadan-Lottick N, Neglia JP. 2014. The contribution of neurocognitive functioning to quality of life after childhood acute lymphoblastic leukemia. *Psychooncology* 23(6):692–699, PMID: 24497266, <https://doi.org/10.1002/pon.3470>.
- Hunger SP, Mullighan CG. 2015. Acute lymphoblastic leukemia in children. *N Engl J Med* 373(16):1541–1552, PMID: 26465987, <https://doi.org/10.1056/NEJMra1400972>.
- Greaves M. 2006. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer* 6(3):193–203, PMID: 16467884, <https://doi.org/10.1038/nrc1816>.
- Greaves MF, Maia AT, Wiemels JL, Ford AM. 2003. Leukemia in twins: lessons in natural history. *Blood* 102(7):2321–2333, PMID: 12791663, <https://doi.org/10.1182/blood-2002-12-3817>.
- Wiemels JL, Cazzaniga G, Daniotti M, Eden OB, Addison GM, Masera G, et al. 1999. Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet* 354(9189):1499–1503, PMID: 10551495, [https://doi.org/10.1016/S0140-6736\(99\)09403-9](https://doi.org/10.1016/S0140-6736(99)09403-9).
- Buffler PA, Kwan ML, Reynolds P, Urayama KY. 2005. Environmental and genetic risk factors for childhood leukemia: appraising the evidence. *Cancer Invest* 23(1):60–75, PMID: 15779869, <https://doi.org/10.1081/CNV-46402>.
- Bailey HD, Armstrong BK, de Klerk NH, Fritschi L, Attia J, Lockwood L, et al. 2010. Exposure to diagnostic radiological procedures and the risk of childhood acute lymphoblastic leukemia. *Cancer Epidemiol Biomarkers Prev* 19(11):2897–2909, PMID: 20861400, <https://doi.org/10.1158/1055-9965.EPI-10-0542>.
- Bartley K, Metayer C, Selvin S, Ducore J, Buffler P. 2010. Diagnostic X-rays and risk of childhood leukaemia. *Int J Epidemiol* 39(6):1628–1637, PMID: 20889538, <https://doi.org/10.1093/ije/dyq162>.
- Tong J, Qin L, Cao Y, Li J, Zhang J, Nie J, et al. 2012. Environmental radon exposure and childhood leukemia. *J Toxicol Environ Health B Crit Rev* 15(5):332–347, PMID: 22852813, <https://doi.org/10.1080/10937404.2012.689555>.
- Ghosh JK, Heck JE, Cockburn M, Su J, Jerrett M, Ritz B. 2013. Prenatal exposure to traffic-related air pollution and risk of early childhood cancers. *Am J Epidemiol* 178(8):1233–1239, PMID: 23989198, <https://doi.org/10.1093/aje/kwt129>.
- Filippini T, Hatch EE, Rothman KJ, Heck JE, Park AS, Crippa A, et al. 2019. Association between outdoor air pollution and childhood leukemia: a systematic review and dose–response meta-analysis. *Environ Health Perspect* 127(4):046002, PMID: 31017485, <https://doi.org/10.1289/EHP4381>.
- Heck JE, Park AS, Qiu J, Cockburn M, Ritz B. 2014. Risk of leukemia in relation to exposure to ambient air toxics in pregnancy and early childhood. *Int J Hyg Environ Health* 217(6):662–668, PMID: 24472648, <https://doi.org/10.1016/j.ijheh.2013.12.003>.
- Heck JE, Wu J, Lombardi C, Qiu J, Meyers TJ, Wilhelm M, et al. 2013. Childhood cancer and traffic-related air pollution exposure in pregnancy and early life. *Environ Health Perspect* 121(11–12):1385–1391, PMID: 24021746, <https://doi.org/10.1289/ehp.1306761>.
- Symanski E, Tee Lewis PG, Chen TY, Chan W, Lai D, Ma X. 2016. Air toxics and early childhood acute lymphocytic leukemia in Texas, a population based case control study. *Environ Health* 15(1):70, PMID: 27301866, <https://doi.org/10.1186/s12940-016-0154-8>.
- Bailey HD, Infante-Rivard C, Metayer C, Clavel J, Lightfoot T, Kaatsch P, et al. 2015. Home pesticide exposures and risk of childhood leukemia: findings from the childhood leukemia international consortium. *Int J Cancer* 137(11):2644–2663, PMID: 26061779, <https://doi.org/10.1002/ijc.29631>.
- Park AS, Ritz B, Yu F, Cockburn M, Heck JE. 2020. Prenatal pesticide exposure and childhood leukemia – a California statewide case-control study. *Int J Hyg Environ Health* 226:113486, PMID: 32087503, <https://doi.org/10.1016/j.ijheh.2020.113486>.
- Wigle DT, Turner MC, Krewski D. 2009. A systematic review and meta-analysis of childhood leukemia and parental occupational pesticide exposure. *Environ Health Perspect* 117(10):1505–1513, PMID: 20019898, <https://doi.org/10.1289/ehp.0900582>.
- Gunier RB, Kang A, Hammond SK, Reinier K, Lea CS, Chang JS, et al. 2017. A task-based assessment of parental occupational exposure to pesticides and childhood acute lymphoblastic leukemia. *Environ Res* 156:57–62, PMID: 28319818, <https://doi.org/10.1016/j.envres.2017.03.001>.
- Ma X, Buffler PA, Gunier RB, Dahl G, Smith MT, Reinier K, et al. 2002. Critical windows of exposure to household pesticides and risk of childhood leukemia. *Environ Health Perspect* 110(9):955–960, PMID: 12204832, <https://doi.org/10.1289/ehp.02110955>.
- Ward MH, Colt JS, Deziel NC, Whitehead TP, Reynolds P, Gunier RB, et al. 2014. Residential levels of polybrominated diphenyl ethers and risk of childhood acute lymphoblastic leukemia in California. *Environ Health Perspect* 122(10):1110–1116, PMID: 24911217, <https://doi.org/10.1289/ehp.1307602>.
- Smith MT. 1996. The mechanism of benzene-induced leukemia: a hypothesis and speculations on the causes of leukemia. *Environ Health Perspect* 104(suppl 6):1219–1225, PMID: 9118896, <https://doi.org/10.1289/ehp.961041219>.
- Heck JE, He D, Contreras ZA, Ritz B, Olsen J, Hansen J. 2019. Parental occupational exposure to benzene and the risk of childhood and adolescent acute lymphoblastic leukaemia: a population-based study. *Occup Environ Med* 76(8):527–529, PMID: 31138675, <https://doi.org/10.1136/oemed-2019-105738>.
- Snyder R. 2012. Leukemia and benzene. *Int J Environ Res Public Health* 9(8):2875–2893, PMID: 23066403, <https://doi.org/10.3390/ijerph9082875>.
- Savitz DA, Andrews KW. 1997. Review of epidemiologic evidence on benzene and lymphatic and hematopoietic cancers. *Am J Ind Med* 31(3):287–295, PMID: 9055951, [https://doi.org/10.1002/\(sici\)1097-0274\(199703\)31:3<287::aid-ajim4>3.0.co;2-v](https://doi.org/10.1002/(sici)1097-0274(199703)31:3<287::aid-ajim4>3.0.co;2-v).
- Zhou Y, Zhang S, Li Z, Zhu J, Bi Y, Bai Y, et al. 2014. Maternal benzene exposure during pregnancy and risk of childhood acute lymphoblastic leukemia: a meta-analysis of epidemiologic studies. *PLoS One* 9(10):e110466, PMID: 25333868, <https://doi.org/10.1371/journal.pone.0110466>.
- Deziel NC, Clark CJ, Casey JA, Bell ML, Plata DL, Saiers JE. 2022. Assessing exposure to unconventional oil and gas development: strengths, challenges, and implications for epidemiologic research. *Curr Environ Health Rep*, PMID: 35522388, <https://doi.org/10.1007/s40572-022-00358-4>.
- U.S. EPA (United States Environmental Protection Agency). 2021. The Process of Unconventional Natural Gas Production. <https://www.epa.gov/uog/process-unconventional-natural-gas-production> [accessed 15 January].
- Kondash AJ, Albright E, Vengosh A. 2017. Quantity of flowback and produced waters from unconventional oil and gas exploration. *Sci Total Environ* 574:314–321, PMID: 27639468, <https://doi.org/10.1016/j.scitotenv.2016.09.069>.
- Scanlon BR, Reedy RC, Male F, Walsh M. 2017. Water issues related to transitioning from conventional to unconventional oil production in the Permian Basin. *Environ Sci Technol* 51(18):10903–10912, PMID: 28876906, <https://doi.org/10.1021/acs.est.7b02185>.

40. Cozzarelli IM, Skalak KJ, Kent DB, Engle MA, Benthem A, Mumford AC, et al. 2017. Environmental signatures and effects of an oil and gas wastewater spill in the Williston Basin, North Dakota. *Sci Total Environ* 579:1781–1793, PMID: [27939081](https://doi.org/10.1016/j.scitotenv.2016.11.157), <https://doi.org/10.1016/j.scitotenv.2016.11.157>.
41. Drollette BD, Hoelzer K, Warner NR, Darrah TH, Karatum O, O'Connor MP, et al. 2015. Elevated levels of diesel range organic compounds in groundwater near Marcellus gas operations are derived from surface activities. *Proc Natl Acad Sci USA* 112(43):13184–13189, PMID: [26460018](https://doi.org/10.1073/pnas.1511474112), <https://doi.org/10.1073/pnas.1511474112>.
42. Maloney KO, Baruch-Mordo S, Patterson LA, Nicot J-P, Entrekin SA, Fargione JE, et al. 2017. Unconventional oil and gas spills: materials, volumes, and risks to surface waters in four states of the U.S. *Sci Total Environ* 581–582:369–377, PMID: [28043701](https://doi.org/10.1016/j.scitotenv.2016.12.142), <https://doi.org/10.1016/j.scitotenv.2016.12.142>.
43. Patterson LA, Konschnik KE, Wiseman H, Fargione J, Maloney KO, Kiesecker J, et al. 2017. Unconventional oil and gas spills: risks, mitigation priorities, and state reporting requirements. *Environ Sci Technol* 51(5):2563–2573, PMID: [28220696](https://doi.org/10.1021/acs.est.6b05749), <https://doi.org/10.1021/acs.est.6b05749>.
44. DiGiulio DC, Jackson RB. 2016. Impact to underground sources of drinking water and domestic wells from production well stimulation and completion practices in the Pavillion, Wyoming, field. *Environ Sci Technol* 50(8):4524–4536, PMID: [27022977](https://doi.org/10.1021/acs.est.5b04970), <https://doi.org/10.1021/acs.est.5b04970>.
45. Vengosh A, Jackson RB, Warner N, Darrah TH, Kondash A. 2014. A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States. *Environ Sci Technol* 48(15):8334–8348, PMID: [24606408](https://doi.org/10.1021/es405118y), <https://doi.org/10.1021/es405118y>.
46. Orem W, Varonka M, Crosby L, Haase K, Loftin K, Hladik M, et al. 2017. Organic geochemistry and toxicology of a stream impacted by unconventional oil and gas wastewater disposal operations. *Appl Geochem* 80:155–167, <https://doi.org/10.1016/j.apgeochem.2017.02.016>.
47. Elliott EG, Trinh P, Ma X, Leaderer BP, Ward MH, Deziel NC. 2017. Unconventional oil and gas development and risk of childhood leukemia: assessing the evidence. *Sci Total Environ* 576:138–147, PMID: [27783932](https://doi.org/10.1016/j.scitotenv.2016.10.072), <https://doi.org/10.1016/j.scitotenv.2016.10.072>.
48. U.S. EPA. 2016. Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States (Final Report). EPA/600/R-16/236F, 2016. Washington, DC: U.S. Environmental Protection Agency.
49. Sumner AJ, Plata Desiree L. 2018. Exploring the hydraulic fracturing parameter space: a novel high-pressure, high-throughput reactor system for investigating subsurface chemical transformations. *Environ Sci Process Impacts* 20(2):318–331, PMID: [29313541](https://doi.org/10.1039/c7em00470b), <https://doi.org/10.1039/c7em00470b>.
50. Allshouse WB, McKenzie LM, Barton K, Brindley S, Adgate JL. 2019. Community noise and air pollution exposure during the development of a multi-well oil and gas pad. *Environ Sci Technol* 53(12):7126–7135, PMID: [31136715](https://doi.org/10.1021/acs.est.9b00052), <https://doi.org/10.1021/acs.est.9b00052>.
51. Adgate JL, Goldstein BD, McKenzie LM. 2014. Potential public health hazards, exposures and health effects from unconventional natural gas development. *Environ Sci Technol* 48(15):8307–8320, PMID: [24564405](https://doi.org/10.1021/es404621d), <https://doi.org/10.1021/es404621d>.
52. Macey GP, Breech R, Cherniak M, Cox C, Larson D, Thomas D, et al. 2014. Air concentrations of volatile compounds near oil and gas production: a community-based exploratory study. *Environ Health* 13(1):82, PMID: [25355625](https://doi.org/10.1186/1476-069X-13-82), <https://doi.org/10.1186/1476-069X-13-82>.
53. McKenzie LM, Witter RZ, Newman LS, Adgate JL. 2012. Human health risk assessment of air emissions from development of unconventional natural gas resources. *Sci Total Environ* 424:79–87, PMID: [22444058](https://doi.org/10.1016/j.scitotenv.2012.02.018), <https://doi.org/10.1016/j.scitotenv.2012.02.018>.
54. Goetz JD, Floerchinger C, Fortner EC, Wormhoudt J, Massoli P, Knighton WB, et al. 2015. Atmospheric emission characterization of Marcellus shale natural gas development sites. *Environ Sci Technol* 49(11):7012–7020, PMID: [25897974](https://doi.org/10.1021/acs.est.5b00452), <https://doi.org/10.1021/acs.est.5b00452>.
55. Li L, Blomberg AJ, Spengler JD, Coull BA, Schwartz JD, Koutrakis P. 2020. Unconventional oil and gas development and ambient particle radioactivity. *Nat Commun* 11(1):5002, PMID: [33051463](https://doi.org/10.1038/s41467-020-18226-w), <https://doi.org/10.1038/s41467-020-18226-w>.
56. Casey JA, Ogburn EL, Rasmussen SG, Irving JK, Pollak J, Locke PA, et al. 2015. Predictors of indoor radon concentrations in Pennsylvania, 1989–2013. *Environ Health Perspect* 123(11):1130–1137, PMID: [25856050](https://doi.org/10.1289/ehp.1409014), <https://doi.org/10.1289/ehp.1409014>.
57. Xu Y, Sajja M, Kumar A. 2019. Impact of the hydraulic fracturing on indoor radon concentrations in Ohio: a multilevel modeling approach. *Front Public Health* 7(7):76, PMID: [31024877](https://doi.org/10.3389/fpubh.2019.00076), <https://doi.org/10.3389/fpubh.2019.00076>.
58. Eitrehm ES, May D, Forbes TZ, Nelson AW. 2016. Disequilibrium of naturally occurring radioactive materials (NORM) in drill cuttings from a horizontal drilling operation. *Environ Sci Technol Lett* 3(12):425–429, <https://doi.org/10.1021/acs.estlett.6b00439>.
59. Rich AL, Crosby EC. 2013. Analysis of reserve pit sludge from unconventional natural gas hydraulic fracturing and drilling operations for the presence of technologically enhanced naturally occurring radioactive material (TENORM). *New Solut* 23(1):117–135, PMID: [23552651](https://doi.org/10.2190/NS.23.1.h), <https://doi.org/10.2190/NS.23.1.h>.
60. Fryzek J, Pastula S, Jiang X, Garabrant DH. 2013. Childhood cancer incidence in Pennsylvania counties in relation to living in counties with hydraulic fracturing sites. *J Occup Environ Med* 55(7):796–801, PMID: [23836020](https://doi.org/10.1097/JOM.0b013e318289ee02), <https://doi.org/10.1097/JOM.0b013e318289ee02>.
61. Goldstein BD, Malone S. 2013. Obfuscation does not provide comfort: response to the article by Fryzek et al on hydraulic fracturing and childhood cancer. *J Occup Environ Med* 55(11):1376–1378, PMID: [24202245](https://doi.org/10.1097/JOM.0000000000000014), <https://doi.org/10.1097/JOM.0000000000000014>.
62. McKenzie LM, Allshouse WB, Byers TE, Bedrick EJ, Serdar B, Adgate JL. 2017. Childhood hematologic cancer and residential proximity to oil and gas development. *PLoS One* 12(2):e0170423, PMID: [28199334](https://doi.org/10.1371/journal.pone.0170423), <https://doi.org/10.1371/journal.pone.0170423>.
63. Pennsylvania Pressroom. 2020. Wolf Administration Awards \$2.5 Million Contract to University of Pittsburgh to Research Health Effects of Hydraulic Fracturing in Pennsylvania. <https://www.media.pa.gov/pages/health-details.aspx?newsid=1215> [accessed 20 August 2021].
64. Soriano MA, Siegel HG, Johnson NP, Gutches KM, Xiong B, Li Y, et al. 2021. Assessment of groundwater well vulnerability to contamination through physics-informed machine learning. *Environ Res Lett* 16(8):084013, <https://doi.org/10.1088/1748-9326/ac10e0>.
65. Clark CJ, Xiong B, Soriano MA, Gutches K, Siegel HG, Ryan EC, et al. 2022. Assessing unconventional oil and gas exposure in the Appalachian Basin: comparison of exposure surrogates and residential drinking water measurements. *Environ Sci Technol* 56(2):1091–1103, PMID: [34982938](https://doi.org/10.1021/acs.est.1c05081), <https://doi.org/10.1021/acs.est.1c05081>.
66. Pennsylvania Department of Environmental Protection. Oil and Gas Reports. <https://www.dep.pa.gov/DataandTools/Reports/Oil%20and%20Gas%20Reports/Pages/default.aspx> [accessed 10 January 2021].
67. Brantley SL, Yoxheimer D, Arjmand S, Grieve P, Vidic R, Pollak J, et al. 2014. Water resource impacts during unconventional shale gas development: the Pennsylvania experience. *Int J Coal Geol* 126:140–156, <https://doi.org/10.1016/j.coal.2013.12.017>.
68. Jasechko S, Perrone D. 2017. Hydraulic fracturing near domestic groundwater wells. *Proc Natl Acad Sci USA* 114(50):13138–13143, PMID: [29180405](https://doi.org/10.1073/pnas.1701682114), <https://doi.org/10.1073/pnas.1701682114>.
69. Barrington-Trimis JL, Cockburn M, Metayer C, Gauderman WJ, Wiemels J, McKean-Cowdin R. 2017. Trends in childhood leukemia incidence over two decades from 1992 to 2013. *Int J Cancer* 140(5):1000–1008, PMID: [27778348](https://doi.org/10.1002/ijc.30487), <https://doi.org/10.1002/ijc.30487>.
70. Doerrenberg M, Kloetgen A, Hezaveh K, Wössmann W, Bleckmann K, Stanulla M, et al. 2017. T-cell acute lymphoblastic leukemia in infants has distinct genetic and epigenetic features compared to childhood cases. *Genes Chromosomes Cancer* 56(2):159–167, PMID: [27717083](https://doi.org/10.1002/gcc.22423), <https://doi.org/10.1002/gcc.22423>.
71. Wiemels J. 2008. Chromosomal translocations in childhood leukemia: natural history, mechanisms, and epidemiology. *J Natl Cancer Inst Monogr* 2008(39):87–90, PMID: [18648011](https://doi.org/10.1093/jncimonographs/ign006), <https://doi.org/10.1093/jncimonographs/ign006>.
72. Howard J. 2012. Minimum Latency and Types or Categories of Cancer. Atlanta, GA: Centers for Disease Control and Prevention, WTC Health Program.
73. Hjalgrim LL, Westergaard T, Rostgaard K, Schmiegelow K, Melbye M, Hjalgrim H, et al. 2003. Birth weight as a risk factor for childhood leukemia: a meta-analysis of 18 epidemiologic studies. *Am J Epidemiol* 158(8):724–735, PMID: [14561661](https://doi.org/10.1093/aje/kwg210), <https://doi.org/10.1093/aje/kwg210>.
74. Cnattingius S, Zack MM, Ekbohm A, Gunnarskog J, Kreuger A, Linet M, et al. 1995. Prenatal and neonatal risk factors for childhood lymphatic leukemia. *J Natl Cancer Inst* 87(12):908–914, PMID: [7666480](https://doi.org/10.1093/jnci/87.12.908), <https://doi.org/10.1093/jnci/87.12.908>.
75. Wang R, Wiemels JL, Metayer C, Morimoto L, Francis SS, Kadan-Lottick N, et al. 2017. Cesarean section and risk of childhood acute lymphoblastic leukemia in a population-based, record-linkage study in California. *Am J Epidemiol* 185(2):96–105, PMID: [27986703](https://doi.org/10.1093/aje/kww153), <https://doi.org/10.1093/aje/kww153>.
76. Wang C, Gomez-Velez JD, Wilson JL. 2018. The importance of capturing topographic features for modeling groundwater flow and transport in mountainous watersheds. *Water Resour Res* 54(12):10313–10338, <https://doi.org/10.1029/2018WR023863>.
77. Soriano MA Jr., Siegel HG, Gutches KM, Clark CJ, Li Y, Xiong B, et al. 2020. Evaluating domestic well vulnerability to contamination from unconventional oil and gas development sites. *Water Resour Res* 56(10):e2020WR028005, <https://doi.org/10.1029/2020WR028005>.
78. Goodman PS, Galatioto F, Thorpe N, Namdeo AK, Davies RJ, Bird RN. 2016. Investigating the traffic-related environmental impacts of hydraulic-fracturing

- (fracking) operations. *Environ Int* 89–90:248–260, PMID: [26922565](https://doi.org/10.1016/j.envint.2016.02.002), <https://doi.org/10.1016/j.envint.2016.02.002>.
79. Llewellyn GT, Dorman F, Westland JL, Yoxtheimer D, Grieve P, Sowers T, et al. 2015. Evaluating a groundwater supply contamination incident attributed to Marcellus shale gas development. *Proc Natl Acad Sci USA* 112(20):6325–6330, PMID: [25941400](https://doi.org/10.1073/pnas.1420279112), <https://doi.org/10.1073/pnas.1420279112>.
 80. Henri CV, Harter T. 2019. Stochastic assessment of nonpoint source contamination: joint impact of aquifer heterogeneity and well characteristics on management metrics. *Water Resour Res* 55(8):6773–6794, <https://doi.org/10.1029/2018WR024230>.
 81. Rice AK, McCray JE, Singha K. 2021. Numerical investigation of wellbore methane leakage from a dual-porosity reservoir and subsequent transport in groundwater. *Water Resour Res* 57(2):e2019WR026991, <https://doi.org/10.1029/2019WR026991>.
 82. Bell ML, Banerjee G, Pereira G. 2018. Residential mobility of pregnant women and implications for assessment of spatially-varying environmental exposures. *J Expo Sci Environ Epidemiol* 28(5):470–480, PMID: [29511287](https://doi.org/10.1038/s41370-018-0026-0), <https://doi.org/10.1038/s41370-018-0026-0>.
 83. Bell ML, Belanger K. 2012. Review of research on residential mobility during pregnancy: consequences for assessment of prenatal environmental exposures. *J Expo Sci Environ Epidemiol* 22(5):429–438, PMID: [22617723](https://doi.org/10.1038/jes.2012.42), <https://doi.org/10.1038/jes.2012.42>.
 84. Hodgson S, Lurz PW, Shirley MD, Bythell M, Rankin J. 2015. Exposure misclassification due to residential mobility during pregnancy. *Int J Hyg Environ Health* 218(4):414–421, PMID: [25845985](https://doi.org/10.1016/j.ijheh.2015.03.007), <https://doi.org/10.1016/j.ijheh.2015.03.007>.
 85. Ling C, Heck JE, Cockburn M, Liew Z, Marcotte E, Ritz B. 2019. Residential mobility in early childhood and the impact on misclassification in pesticide exposures. *Environ Res* 173:212–220, PMID: [30928851](https://doi.org/10.1016/j.envres.2019.03.039), <https://doi.org/10.1016/j.envres.2019.03.039>.
 86. Kumar A, Vashist M, Rathee R. 2014. Maternal factors and risk of childhood leukemia. *Asian Pac J Cancer Prev* 15(2):781–784, PMID: [24568495](https://doi.org/10.7314/apjcp.2014.15.2.781), <https://doi.org/10.7314/apjcp.2014.15.2.781>.
 87. Belson M, Kingsley B, Holmes A. 2007. Risk factors for acute leukemia in children: a review. *Environ Health Perspect* 115(1):138–145, PMID: [17366834](https://doi.org/10.1289/ehp.9023), <https://doi.org/10.1289/ehp.9023>.
 88. Wiemels J. 2012. Perspectives on the causes of childhood leukemia. *Chem Biol Interact* 196(3):59–67, PMID: [22326931](https://doi.org/10.1016/j.cbi.2012.01.007), <https://doi.org/10.1016/j.cbi.2012.01.007>.
 89. U.S. EPA. RSIG-Related Downloadable Data Files: Fused Air Quality Surface Using Downscaling (FAQSD) Files. <https://www.epa.gov/hesc/rsig-related-downloadable-data-files> [accessed 24 August 2021].
 90. United States Department of Agriculture. National Agricultural Statistics Service CropScape. <https://nassgeodata.gmu.edu/CropScape/> [accessed 21 July 2021].
 91. Reynolds P, Von Behren J, Gunier RB, Goldberg DE, Harnly M, Hertz A. 2005. Agricultural pesticide use and childhood cancer in California. *Epidemiology* 16(1):93–100, PMID: [15613951](https://doi.org/10.1093/epi/kh111).
 92. United States Census Bureau. <https://data.census.gov/cedsci/> [accessed 21 August 2021].
 93. ATSDR (Agency for Toxic Substances and Disease Registry). The Social Vulnerability Index. <https://www.atsdr.cdc.gov/placeandhealth/svi/index.html> [accessed 21 August 2021].
 94. Bailey HD, Armstrong BK, de Klerk NH, Fritschi L, Attia J, Scott RJ, et al. 2011. Exposure to professional pest control treatments and the risk of childhood acute lymphoblastic leukemia. *Int J Cancer* 129(7):1678–1688, PMID: [21080443](https://doi.org/10.1002/ijc.25769), <https://doi.org/10.1002/ijc.25769>.
 95. Greaves M. 2018. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nat Rev Cancer* 18(8):471–484, PMID: [29784935](https://doi.org/10.1038/s41568-018-0015-6), <https://doi.org/10.1038/s41568-018-0015-6>.
 96. Elliott EG, Ma X, Leaderer BP, McKay LA, Pedersen CJ, Wang C, et al. 2018. A community-based evaluation of proximity to unconventional oil and gas wells, drinking water contaminants, and health symptoms in Ohio. *Environ Res* 167:550–557, PMID: [30145431](https://doi.org/10.1016/j.envres.2018.08.022), <https://doi.org/10.1016/j.envres.2018.08.022>.
 97. McMahon PB, Lindsey BD, Conlon MD, Hunt AG, Belitz K, Jurgens BC, et al. 2019. Hydrocarbons in upland groundwater, Marcellus Shale region, northeastern Pennsylvania and southern New York, U.S.A. *Environ Sci Technol* 53(14):8027–8035, PMID: [31246428](https://doi.org/10.1021/acs.est.9b01440), <https://doi.org/10.1021/acs.est.9b01440>.
 98. Wen T, Niu X, Gonzales M, Zheng G, Li Z, Brantley SL. 2018. Big groundwater data sets reveal possible rare contamination amid otherwise improved water quality for some analytes in a region of Marcellus shale development. *Environ Sci Technol* 52(12):7149–7159, PMID: [29783843](https://doi.org/10.1021/acs.est.8b01123), <https://doi.org/10.1021/acs.est.8b01123>.
 99. Xiong B, Soriano MA, Gutches KM, Hoffman N, Clark CJ, Siegel HG, et al. 2022. Groundwaters in Northeastern Pennsylvania near intense hydraulic fracturing activities exhibit few organic chemical impacts. *Environ Sci: Processes Impacts* 24(2):252–264, PMID: [35018906](https://doi.org/10.1039/D1EM00124H), <https://doi.org/10.1039/D1EM00124H>.
 100. Caron-Bestaudoine É, Valter N, Chevrier J, Ayotte P, Frohlich K, Verner M-A. 2018. Gestational exposure to volatile organic compounds (VOCs) in Northeastern British Columbia, Canada: a pilot study. *Environ Int* 110:131–138, PMID: [29122312](https://doi.org/10.1016/j.envint.2017.10.022), <https://doi.org/10.1016/j.envint.2017.10.022>.
 101. Cushing LJ, Vavra-Musser K, Chau K, Franklin M, Johnston JE. 2020. Flaring from unconventional oil and gas development and birth outcomes in the Eagle Ford Shale in South Texas. *Environ Health Perspect* 128(7):77003, PMID: [32673511](https://doi.org/10.1289/EHP6394), <https://doi.org/10.1289/EHP6394>.
 102. Casey JA, Elser H, Goldman-Mellor S, Catalano R. 2019. Increased motor vehicle crashes following induced earthquakes in Oklahoma, USA. *Sci Total Environ* 650(pt 2):2974–2979, PMID: [30373073](https://doi.org/10.1016/j.scitotenv.2018.10.043), <https://doi.org/10.1016/j.scitotenv.2018.10.043>.
 103. Allshouse WB, Adgate JL, Blair BD, McKenzie LM. 2017. Spatiotemporal industrial activity model for estimating the intensity of oil and gas operations in Colorado. *Environ Sci Technol* 51(17):10243–10250, PMID: [28715172](https://doi.org/10.1021/acs.est.7b02084), <https://doi.org/10.1021/acs.est.7b02084>.
 104. Casey JA, Savitz DA, Rasmussen SG, Ogburn EL, Pollak J, Mercer DG, et al. 2016. Unconventional natural gas development and birth outcomes in Pennsylvania. *Epidemiology* 27(2):163–172, PMID: [26426945](https://doi.org/10.1097/EDE.0000000000000387), <https://doi.org/10.1097/EDE.0000000000000387>.
 105. Stacy SL, Brink LL, Larkin JC, Sadovsky Y, Goldstein BD, Pitt BR, et al. 2015. Perinatal outcomes and unconventional natural gas operations in Southwest Pennsylvania. *PLoS One* 10(6):e0126425, PMID: [26039051](https://doi.org/10.1371/journal.pone.0126425), <https://doi.org/10.1371/journal.pone.0126425>.
 106. Rogers JD, Burke TL, Osborn SG, Ryan JN. 2015. A framework for identifying organic compounds of concern in hydraulic fracturing fluids based on their mobility and persistence in groundwater. *Environ Sci Technol Lett* 2(6):158–164, <https://doi.org/10.1021/acs.estlett.5b00090>.
 107. Osborn SG, Vengosh A, Warner NR, Jackson RB. 2011. Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing. *Proc Natl Acad Sci USA* 108(20):8172–8176, PMID: [21555547](https://doi.org/10.1073/pnas.1100682108), <https://doi.org/10.1073/pnas.1100682108>.
 108. Jackson RE, Gorody AW, Mayer B, Roy JW, Ryan MC, Van Stempvoort DR. 2013. Groundwater protection and unconventional gas extraction: the critical need for field-based hydrogeological research. *Ground Water* 51(4):488–510, PMID: [23745972](https://doi.org/10.1111/gwat.12074), <https://doi.org/10.1111/gwat.12074>.
 109. Deziel NC. 2021. Invited perspective: oil and gas development and adverse birth outcomes: what more do we need to know? *Environ Health Perspect* 129(7):71301, PMID: [34287014](https://doi.org/10.1289/EHP9715), <https://doi.org/10.1289/EHP9715>.
 110. Shonkoff SBC, Morello-Frosch R, Casey JA, Deziel NC, Digiulio DC, Foster S, et al. 2021. RE: Response to CalGEM Questions for the California Oil and Gas Public Health Rulemaking Scientific Advisory Panel. https://www.conservation.ca.gov/calgem/Documents/public-health/Public%20Health%20Panel%20Responses_FINAL%20ADA.pdf [accessed 1 February 2022].
 111. Lewis C, Greiner LH, Brown DR. 2018. Setback distances for unconventional oil and gas development: Delphi study results. *PLoS One* 13(8):e0202462, PMID: [30114206](https://doi.org/10.1371/journal.pone.0202462), <https://doi.org/10.1371/journal.pone.0202462>.
 112. Banan Z, Gernand JM. 2018. Evaluation of gas well setback policy in the Marcellus Shale region of Pennsylvania in relation to emissions of fine particulate matter. *J Air Waste Manag Assoc* 68(9):988–1000, PMID: [29667538](https://doi.org/10.1080/10962247.2018.1462866), <https://doi.org/10.1080/10962247.2018.1462866>.
 113. Environmental Health Project. 2022. Position Statement on Setback Distances: EHP recommends increased setback distances from shale gas operations to protect public health and safety. <https://www.environmentalhealthproject.org/sites/default/files/assets/attachments/ehp-position-statement-setback-distances.pdf> [accessed 25 January 2020].
 114. Colorado Oil & Gas Conservation Commission. 2020. Colorado Oil & Gas Conservation Commission Unanimously Adopts SB 19-181 New Mission Change Rules, Alternative Location Analysis and Cumulative Impacts. <https://dnr.colorado.gov/press-release/colorado-oil-gas-conservation-commission-unanimously-adopts-sb-19-181-new-mission> [accessed 1 February 2022].
 115. Geologic Energy Management Division (CalGEM). 2021. Draft Rule for Protection of Communities and Workers from Health and Safety Impacts from Oil and Gas Production Operations Pre-Rulemaking Release for Public Review and Consultation. Sacramento, CA: California Department of Conservation Geologic Energy Management Division. <https://www.conservation.ca.gov/calgem/Documents/public-health/PHRM%20Draft%20Rule.pdf> [accessed 1 February 2022].
 116. Pennsylvania Department of Environmental Protection. Act 13 of 2012. 2012. <https://www.dep.pa.gov/Business/Energy/OilandGasPrograms/Act13/Pages/default.aspx> [accessed 15 June 2020].
 117. Zahnd WE, James AS, Jenkins WD, Izadi SR, Fogleman AJ, Steward DE, et al. 2018. Rural–urban differences in cancer incidence and trends in the United

- States. *Cancer Epidemiol Biomarkers Prev* 27(11):1265–1274, PMID: [28751476](https://pubmed.ncbi.nlm.nih.gov/28751476/), <https://doi.org/10.1158/1055-9965.EPI-17-0430>.
118. Clark LP, Millet DB, Marshall JD. 2014. National patterns in environmental injustice and inequality: outdoor NO₂ air pollution in the United States. *PLoS One* 9(4):e94431, PMID: [24736569](https://pubmed.ncbi.nlm.nih.gov/24736569/), <https://doi.org/10.1371/journal.pone.0094431>.
119. Brønnum-Hansen H, Bender AM, Andersen ZJ, Sørensen J, Bønløkke JH, Boshuizen H, et al. 2018. Assessment of impact of traffic-related air pollution on morbidity and mortality in Copenhagen municipality and the health gain of reduced exposure. *Environ Int* 121(pt 1):973–980, PMID: [30408890](https://pubmed.ncbi.nlm.nih.gov/30408890/), <https://doi.org/10.1016/j.envint.2018.09.050>.
120. Clune JW, Cravotta CA III. 2019. Drinking Water Health Standards Comparison and Chemical Analysis of Groundwater for 72 Domestic Wells in Bradford County, Pennsylvania, 2016. Reston, VA: U.S. Geological Survey, 76.
121. Angle MP, Jonak J. 2002. Ground Water Pollution Potential of Belmont County, Ohio: Ground Water Pollution Potential Report No. 50. Columbus, OH: Ohio Department of Natural Resources.