APPENDIX AVAILABLE ON
THE HEALTH EFFECTS INSTITUTE–ENERGY WEBSITE

Special Report 1

POTENTIAL HUMAN HEALTH EFFECTS ASSOCIATED WITH UNCONVENTIONAL OIL AND GAS DEVELOPMENT: A SYSTEMATIC REVIEW OF THE EPIDEMIOLOGY LITERATURE

HEI-Energy Research Committee

APPENDIX B. Summaries of Individual Studies and Their Strengths and Limitations

Correspondence may be addressed to Dr. Donna Vorhees, Health Effects Institute–Energy, 75 Federal Street, Suite 1400, Boston, MA 02110; e-mail: *dvorhees@healtheffects.org*.

Although this document was produced with partial funding by the United States Environmental Protection Agency under Contract No. 68HERC19D0010 to the Health Effects Institute–Energy, it has not been subject to the Agency's review and therefore does not necessarily reflect the views of the Agency, and no official endorsement by the Agency should be inferred. Private institutions also provided funding to produce this document; however, it has not been subject to their review and therefore does not necessarily reflect the views of any of the private institutions, and no endorsement by them should be inferred.

APPENDIX B

Summaries of Individual Studies and Their Strengths and Limitations

CONTENTS

MCKENZIE ET AL. 2014

Research Objective. The objective of this study was to examine the association between maternal exposure to natural gas development and birth defects, preterm birth, and fetal growth.

Study Period and Location. This study was conducted in Colorado areas and towns with populations <50,000. Data were collected for the years 1996 to 2009.

Study Population. The study population included all women who delivered live, singleton births within the study area between 1996 and 2009. Deliveries of non-white births were excluded, reaching a final sample size of 124,842. There were 887 cases of congenital heart defects (CHD), 27 of neural tube defects (NTD), and 139 cases of oral clefts.

Outcome Ascertainment. All live birth data, including preterm births and term low birth weight data were obtained from the Colorado Vital Birth Statistics. Cases of CHD, NTD, and oral clefts were identified in the "Colorado Responds to Children with Special Needs" birth registry and matched to birth certificates. In an exploratory analysis, the investigators stratified total CHDs into seven clinical diagnostic groups: conotruncal defects, ventricular septal defects, endocardial cushion and mitral valve defects, pulmonary artery and valve defects, tricuspid valve defects, aortic artery and valve defects, and patent ductus arteriosus.

Exposure Assessment. The investigators built a dataset containing geocoded coordinates of all Colorado gas wells and year of development between 1996 and 2009, with all data obtained from the Colorado Oil and Gas Information System (COGIS). The investigators created an inverse distance weighted (IDW) count of all existing natural wells within a 10-mile radius of the mother's address at the time of delivery. The continuous IDW measure was categorized into tertiles. The referent group was mothers with zero wells within 10 miles of their residence at the time of delivery.

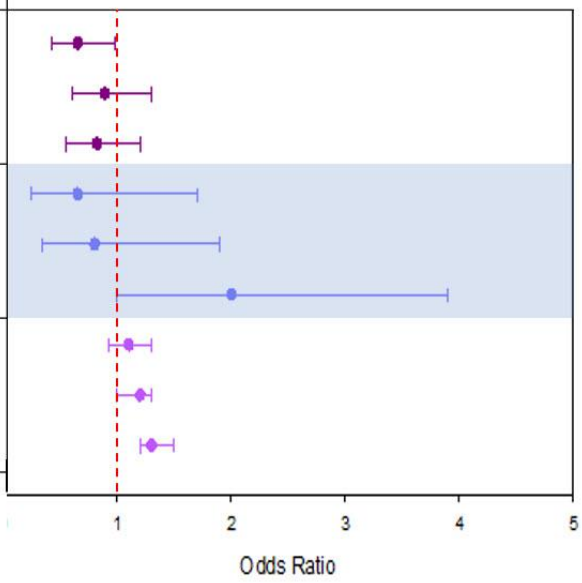
Analytical Methods and Covariate Inclusion. The investigators used logistic regression for all bivariate outcomes and linear regression for term birth weight. They used a Cochran-Armitage test to evaluate trend by levels of IDW exposure. To control potential confounding, the investigators excluded births with CHD, NTD, and oral clefts from pre-term birth and birth weight analytical samples. The investigators adjusted all models for maternal age, education, tobacco use, ethnicity, alcohol use, parity, and infant sex. They also considered elevation of maternal residence and folic acid fortification in a sensitivity analysis.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information). For a comparison of continuous birth weight and preterm birth results across studies, see Figures 4-1 and 4-2, respectively in Section 4.1.2.

McKenzie et al. 2014[1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Neural Tube Defects, Congenital Heart Defects, and Oral Clefts

End Point	Cases	Total *n*	Exposure Group
Oral Clefts	31	19,214	Tertile 1
	41	19,209	Tertile 2
	40	19,793	Tertile 3
Neural Tube Defects	6	19,214	Tertile 1
	7	19,209	Tertile 2
	19	19,793	Tertile 3
Congenital Heart Defects	281	19,214	Tertile 1
	300	19,209	Tertile 2
	355	19,793	Tertile 3

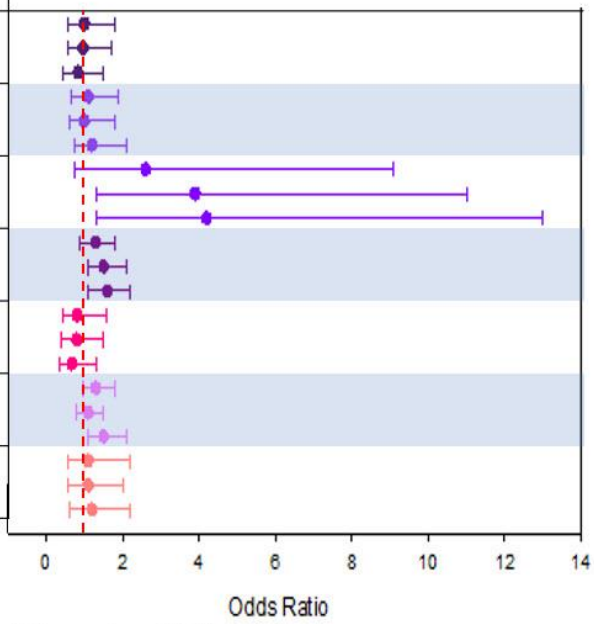


1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure surrogate: IDW calculated within 10 miles of home. Referent: No wells within 10 miles of home.

McKenzie et al. 2014[1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Specific Congenital Heart Defects

End Point	Cases	Total *n*	Exposure Group
Patent Ductus Arteriosus	18	59	Tertile 1
	17		Tertile 2
	15		Tertile 3
Aortic Artery and Valve Defects	22	75	Tertile 1
	21		Tertile 2
	24		Tertile 3
Tricuspid Valve Defects	5	9	Tertile 1
	8		Tertile 2
	8		Tertile 3
Pulmonary Artery and Valve Defects	52	137	Tertile 1
	62		Tertile 2
	66		Tertile 3
Endocardial Cushion and Mitrovalve Defects	14	39	Tertile 1
	12		Tertile 2
	12		Tertile 3
Ventricular Septal Defects	68	210	Tertile 1
	59		Tertile 2
	84		Tertile 3
Conotruncal Defects	14	40	Tertile 1
	13		Tertile 2
	15		Tertile 3



1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure surrogate: IDW calculated within 10 miles of home. Referent: No wells within 10 miles of home.

McKenzie et al. 2014: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Included all birth records from rural areas.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusion based on residence in non-rural areas, race, and non-singleton births.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to retrospective cohort study design.	
	Control group appropriate to address study question	Not applicable to retrospective cohort study design.	
	Same population over study period		No formal assessment of this assumption for the 13-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	All presented characteristics were similar among exposure groups.	
Outcome Assessment	Outcome ascertained using valid and reliable measures	-Ascertained birth defects data from the Colorado Responds to Children with Special Needs (CRCSN) birth registry, identified using ICD-9 codes. -Ascertained birth weight and preterm birth data from Colorado Vital Birth Statistics.	No discussion of quality of CRCSN data.
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Reliability depends on the quality of data available at COGIS.	-No discussion of geocoding methods. -Assumed residential stability throughout study period. -No discussion of quality of COGIS.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.

McKenzie et al. 2014: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Assess exposure in a way that addresses review question.	Incorporated information about proximity to and number of wells (objective to assess exposure to all natural gas wells in study area).	-Did not evaluate whether exposure surrogate represents OGD activities. -Did not differentiate between unconventional and conventional oil and gas wells in analysis but choose study period (1996-2009) to focus on unconventional development.
	Study period sufficient to capture exposure variability	Not applicable to retrospective cohort study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	90% of wells within 7.7 miles of homes.	-Did not test for cut point bias. -No justification for choice of radius within which to calculate exposure surrogate.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells		Not reported.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		No consideration of timing of exposure with respect to gestational period.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data to collect covariate information.	Control of confounding limited by available data on birth records.
	Controlled for baseline conditions	-Controlled for basic maternal sociodemographic characteristics, alcohol consumption, and residential elevation.	-No control of potential confounding by community-level factors correlated with the exposure surrogate or health outcomes. -Lacking detailed information on prenatal care, co-morbidities, SES, and lifestyle factors (except smoking and alcohol).
	Controlled for background exposures	-Limited to rural areas to avoid non-UOGD sources. -Controlled for smoking.	No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells
	Assessed time trends		No.
Analytical Methods	Analytical methods appropriate for study design	Yes.	

McKenzie et al. 2014: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Report measures of precision and variability	-Presented variability of maternal characteristics by exposure group. -95% Confidence intervals for odds ratios and mean differences.	
	Report which statistical tests were used	Cochran-Armitage trend test; $p<0.05$ for significance testing	
	Perform analysis to test sensitivity of results to alternative specifications.	-Tested potential confounding by folic acid fortification. -Restricted analysis to births between 2000 and 2009 to isolate impact of unconventional wells. -Tested sensitivity of alternative buffer distances around maternal residence.	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Appropriate and complete interpretation of results	-Interpreted reported associations appropriately. -Described alternative explanations for findings.	
	Discussion adequately addresses study limitations	Nuanced discussion of limitations in exposure and outcome data, potential for selection bias, limited covariate adjustment in models of low event outcomes.	-No discussion of disparate findings among different distance buffers (2-, 5-, and 10-mile) used to calculate the exposure surrogate. -No discussion of limited covariate adjustment in models with low event counts.

Research Objective. The objective of this study was to examine the association between maternal exposure to unconventional natural gas development and birth outcomes in Southwestern Pennsylvania.

Study Period and Location. Both natural gas well and birth data were collected for the years 2007–2010 for three Southwestern Pennsylvania counties: Butler, Washington, and Westmoreland.

Study Population. The study population included live births (n = 15,451) in Butler, Washington, and Westmoreland counties born between 2007 and 2010 (limited to singleton birth residences within 10 miles of unconventional oil and gas [UOG] wells). Multiple births, records without a geocoded address, and records with missing birth outcome and demographic information were excluded. Investigators also excluded births to mothers with addresses greater than 10 miles from any well.

Outcome Ascertainment. Birth weight, gestational age, maternal risk factors, and child sex data were extracted from birth certificates provided by the Pennsylvania Department of Health. The outcomes of interest included: birth weight (continuous), small for gestational age (SGA), and prematurity.

Exposure Assessment. The investigators created an inverse distance weighted (IDW) count of all wells with both a lateral component and hydraulic fracturing within a 10-mile radius of the mother's address. Investigators identified these wells, spud dates, and whether the well was active at the time of data collection from data provided by the Pennsylvania Department of Environmental Protection (PADEP). Using distances between the mother's address at birth and each well, investigators calculated an IDW metric and used the metric to divide study participants into quartiles of exposure. The referent group included mothers in the first quartile of exposure.

Analytical Methods and Covariate Inclusion. Preliminary analyses were performed using analysis of variance (ANOVA) to compare mean birth weight among groups and a chi-squared test for differences in proportions of SGA and gestational age among groups. Multivariable linear (birth weight) and logistic (SGA and gestational age) regression controlled for child sex, mother's age, educational attainment, pre-pregnancy weight, number of prenatal visits, cigarette smoking during pregnancy, gestational diabetes, Women, Infants, and Children (WIC) assistance, parity, and race. Birth weight models additionally adjusted for gestational age.

Results.[1] The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

[1] Forest plots extracted from study

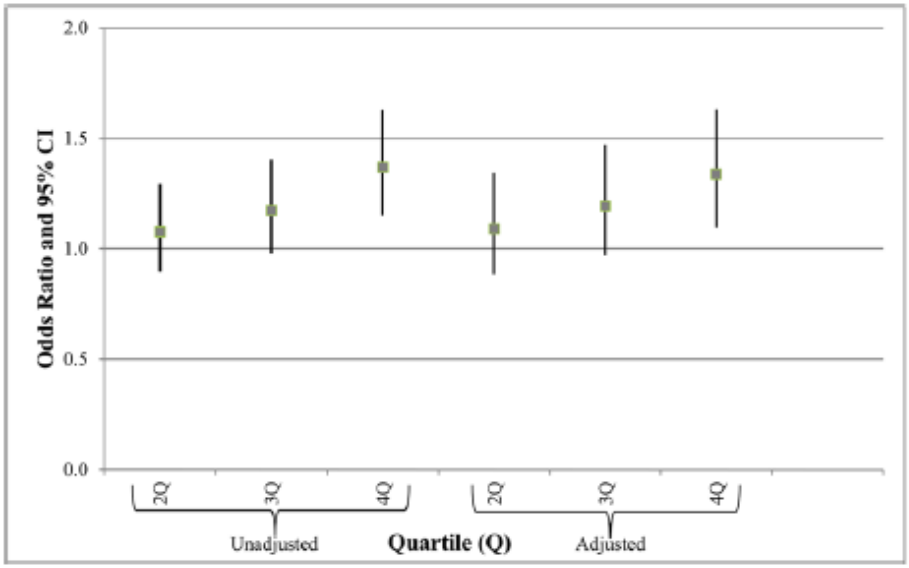


Fig 2. Unadjusted and adjusted odds ratios (OR) and 95% confidence intervals (CI) for small for gestational age (adjusted for mom's age, mom's education, pre-pregnancy weight, gender of infant, prenatal visits, smoking during pregnancy, gestational diabetes, WIC, race, and birth order). Key: Referent (First quartile), <0.87 wells per mile; Second quartile (2Q), 0.87 to 2.59 wells per mile; Third quartile (3Q), 2.60 to 5.99 wells per mile; Fourth quartile (4Q), ≥6.00 wells per mile.

doi:10.1371/journal.pone.0126425.g002

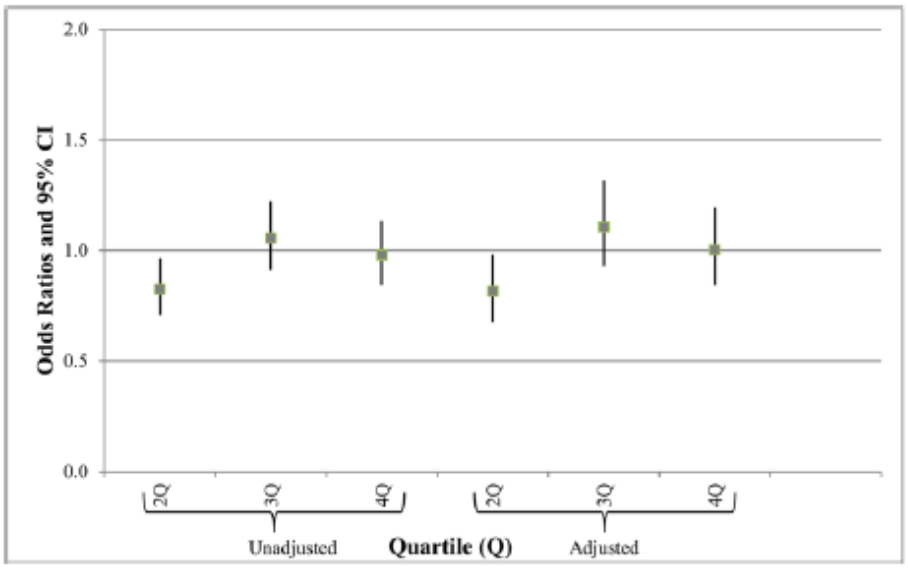


Fig 3. Unadjusted and adjusted odds ratios (OR) and 95% confidence intervals (CI) for prematurity (adjusted for mom's age, mom's education, pre-pregnancy weight, gender of infant, prenatal visits, smoking during pregnancy, gestational diabetes, WIC, race, and birth order). Key: Referent (First quartile), <0.87 wells per mile; Second quartile (2Q), 0.87 to 2.59 wells per mile; Third quartile (3Q), 2.60 to 5.99 wells per mile; Fourth quartile (4Q), ≥6.00 wells per mile.

Stacy et al. 2015: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	-Included all birth records of residence within 10 miles of nearest well. -Presented demographic differences between excluded and included from study populations.	Study sample not representative of general Pennsylvania population with respect to smoking, birth weight, and education.
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on missing data, address not geocoded, and residence >10 miles from closest well.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to retrospective cohort study design.	
	Control group appropriate to address study question	Not applicable to retrospective cohort study design.	
	Same population over study period	Brief study period (2007-2010).	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: prenatal care, gestational diabetes, education, sex, smoking status during pregnancy, parity.	WIC assistance lower in referent compared to other exposure quartiles.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained outcome data from Pennsylvania Department of Health Bureau of Vital Statistics.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Reliability depends on the quality of data available at PADEP.	-No discussion of geocoding methods. -Assumed residential stability throughout study period. -No discussion of quality of data source.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Used two exposure surrogates: 1) proximity to and number of wells, and 2) well density. -Included only UOGD wells.	No evaluation of whether exposure surrogate represents UOGD activities.

Stacy et al. 2015: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Study period sufficient to capture exposure variability	Not applicable to retrospective cohort study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.		-Did not test for cut point bias. -No justification for choice of radius within which to calculate exposure surrogate.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells		Not reported.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Assigned during prenatal period.	Exposure assigned based on year of birth (not date of conception or birth).
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data to collect covariate information.	No control of potential confounding by community-level factors correlated with the exposure surrogate or health outcomes.
	Controlled for baseline conditions	Controlled for basic maternal sociodemographic characteristics (mother's education, WIC), pregnancy risk factors.	No detail control of lifestyle (except smoking), prenatal care, co-morbidities, or detailed SES factors.
	Controlled for background exposures	Controlled for smoking status during pregnancy	No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells
	Assessed time trends		No
Analytical Methods	Analytical methods appropriate for study design	Yes	Did not describe model-building procedure.
	Report measures of precision and variability	-Presented variability of maternal characteristics and outcome measures by exposure group. -95% Confidence intervals for odds ratios.	
	Report which statistical tests were used	-Linear regression (continuous outcomes), logistic regression (binary outcomes). -$p<0.05$ for significance testing.	
	Perform analysis to test sensitivity of results to alternative specifications.		Did not perform sensitivity analyses.
Results and Discussion	All findings reported for analyses described in paper		Did not present well density results.
	Appropriate and complete interpretation of results	Interpreted reported associations appropriately.	No discussion of other potential environmental sources or explanations for observed associations.

Stacy et al. 2015: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Discussion adequately addresses study limitations	-Provided data to address inability to account for residential mobility. -Mentioned lack of fate and transport information included in exposure surrogate, lack of exact birth date.	

Research Objective. The objective of this study was to examine the association between prenatal exposure to unconventional natural gas development and four birth outcomes and high-risk pregnancy.

Study Period and Location. This study was conducted in the 40 central and northeast Pennsylvania counties within the Geisinger Health System catchment area. The study period was January 2009 and January 2013.

Study Population. The study population included mother-child singleton birth pairs, delivered at Geisinger Medical Center and Geisinger Wyoming Valley between 2006 and 2013. The investigators excluded unmatched mother-child records, stillbirths, neonates with serious birth defects, birth weights < 500 g, gestational ages < 22 weeks, non-residents of Pennsylvania, non-geocoded addresses, and births before 2009. The final sample size was 9,384 mothers and 10,496 neonates.

Outcome Ascertainment. The investigators identified births and deliveries using ICD-9 coding from the electronic health records and assessed high-risk pregnancies and four birth outcomes: term birth weight (≥37 week; *n* = 8,839 total live births), preterm birth (<37 week; *n* = 9,848 total live births), low 5-minute Apgar (<7), and SGA.

Exposure Assessment. Well location, spud, production, stimulation dates and drilling depth data from 2005 to 2013 were obtained from the Pennsylvania Department of Environmental Protection and the Pennsylvania Department of Conservation and Natural Resources. Crowd-sourced photos from SkyTruth were used to confirm well location. The investigators developed an activity index for each of the four UOGD development phases. The index incorporated information about distance between wells and maternal residence, dates and durations of development activities and production volume during pregnancy. The four exposure metrics were *z*-transformed and summed to create an aggregate exposure index, categorized into quartiles.

Analytical Methods and Covariate Inclusion. The investigators used multilevel linear (birth weight) and logistic (all other outcomes) models with random intercepts for mother and community of residence. The investigators a priori considered several clinical (e.g., smoking status, parity, and pre-pregnancy BMI), demographic (e.g., race/ethnicity and maternal age) and environmental (e.g., water source, residential greenness, and distance to road) covariates for model inclusion.

Results. The investigators did not provide numeric APGAR scores and small for gestational age results; therefore, results could not be plotted. For a comparison of continuous birth weight and preterm birth results across other studies, see Figures 4-1 and 4-2, respectively, in Section 4.1.2.

Casey et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	-Included all eligible subjects in the Geisinger health system. -Study sample shown to be representative of the general population in a separate study.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on missing data, non-singleton births, serious birth defects, extreme birth weight and gestational age values.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to retrospective cohort study design.	
	Control group appropriate to address study question	Not applicable to retrospective cohort study design.	
	Same population over study period	Reviewed mobility data for full study population and estimated percent expected to move during study period.	Estimated that 20% of study sample expected to move.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: maternal demographic and lifestyle characteristics, residential greenness, prenatal care, season of birth, and infant characteristics.	Proportion of delivery in Geisinger Medical Center and birth year higher in referent group, and proportion of high-risk pregnancy lower in referent compared to exposed groups. Community SES higher in quartile four.
Outcome Assessment	Outcome ascertained using valid and reliable measures	-Used ICD-9 codes from Geisinger Health System medical records. -Reported sources of gestational age calculation.	
	Outcome assessors blinded to exposure status	Obtained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Obtained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	-Reliability depends on the quality of data available at PADEP. -Confirmed well location using crowd sourcing.	-No discussion of geocoding methods. -Assumed residential stability throughout study period using 2013 address, not address at delivery. -No discussion of quality of data source.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.

Casey et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Assess exposure in a way that addresses review question.	-Incorporated information about proximity and number of wells. -Exposure assigned based on date of conception or birth. -Included only UOGD wells. -Assigned at daily-resolution.	No evaluation of whether exposure surrogate represents UOGD activities.
	Study period sufficient to capture exposure variability	Not applicable to retrospective cohort study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.		-Did not test for cut point bias. -Limited spatial overlap between study population and UOGD wells.
	Differentiates among UOGD and its various phases	Differentiated among phases based on recorded dates of spudding, perforation, stimulation, and production.	-Unable to distinguish between phases because of collinearity (aggregated phases to z-score). -Phase duration estimated.
	Differentiates between active and non-active wells	Included active oil and gas wells reported in PADEP database 2005-2013.	Exposure metric may account for wells that became inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Assigned during prenatal period.	
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data to collect detailed covariate information.	
	Controlled for baseline conditions	-Exceptional control of access to prenatal and medical care. -Imprecise measures of individual SES. -Controlled for community-level SES.	No control of co-morbidities or detailed SES factors.
	Controlled for background exposures	Controlled for residential greenness, proximity to major road, water source, and smoking status during pregnancy.	No control of other potential environmental sources, including occupational exposures, industrial sources, or conventional wells
	Assessed time trends	Controlled for year of birth.	
Analytical Methods	Analytical methods appropriate for study design	-Accounted for individual- and community-level correlated standard errors in logistic or linear regression. -Described model-building procedure.	

Casey et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Report measures of precision and variability	-Presented variability of maternal characteristics and outcome measures by exposure group. -95% Confidence intervals for odds ratios.	
	Report which statistical tests were used	-Linear regression (continuous outcomes), logistic regression (binary outcomes). -$p<0.05$ for significance testing.	
	Perform analysis to test sensitivity of results to alternative specifications.	-Conducted sensitivity analysis using a negative exposure control. -Conducted sensitivity analyses restricting to late preterm births, including a measure of pregnancy risk, assessed birth in 2006, before UOGD, applying exposure metrics of later years. and fitting a survival model predicting preterm birth with gestational age for time.	
Results and Discussion	All findings reported for analyses described in paper		Did not report numerical small for gestational age or APGAR results.
	Appropriate and complete interpretation of results	Interpreted reported associations appropriately.	
	Discussion adequately addresses study limitations	-Described alternative explanations for findings. -Provided data to address inability to account for residential mobility. -Discuss some limitations of exposure assessment. -Discuss record reliability.	

Research Objective. The objective of this study was to examine the association between UOGD (defined by the authors as *unconventional natural gas development* [UNGD]) and change in the prevalence of birth defects.

Study Period and Location. Statewide Pennsylvania natural gas and birth data were collected for the years 2003–2012.

Study Population. The study population included all live births in Pennsylvania between 2003 and 2012 (*n* = 1,401,813). Births missing gestational age or ZIP code data were excluded.

Outcome Ascertainment. The investigators ascertained birth defects data from Pennsylvania birth certificates and defined outcomes as binary (yes/no) for structural birth defects, functional/developmental birth defects, and any birth defects.

Exposure Assessment. Two exposure metrics were created using UNGD well data from the Pennsylvania Department of Environmental Protection: earliest spud date within a ZIP code and well density (number of wells per square kilometer) in each ZIP code. Exposure was assigned based on date of conception, which was estimated using gestational age. Exposure metrics were assigned based on the mother's ZIP code at delivery.

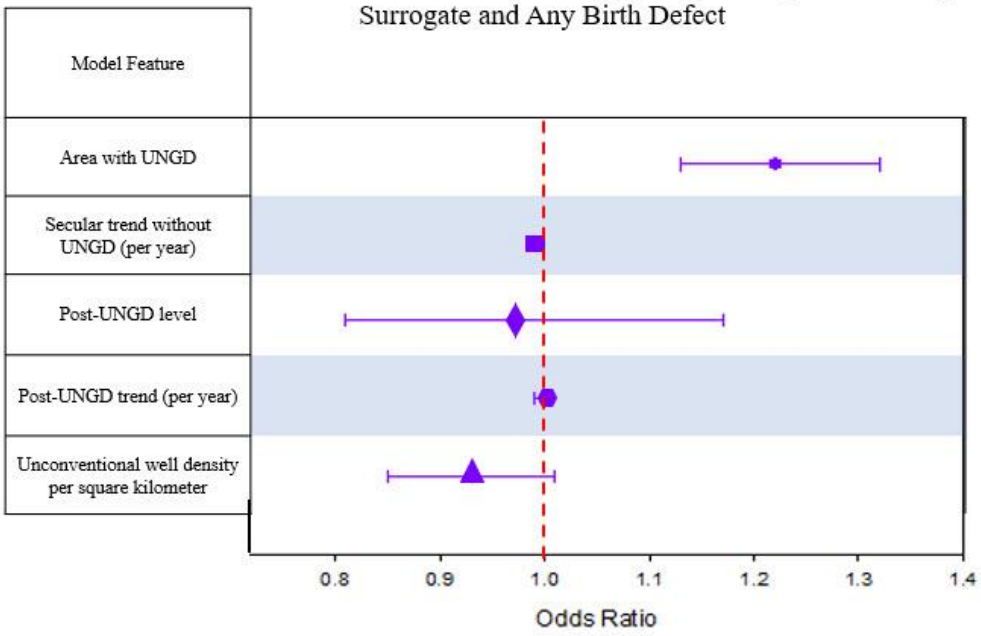
Analytical Methods and Covariate Inclusion. The investigators performed a segmented regression analysis, controlling for maternal characteristics and four derived variables necessary for this analytical method: (1) a continuous time variable indicating the number of months "without UNGD," where a birth in the first month of study (January 2003) would have a value of "1," and a birth in the second month of study (February 2003) would have a value of "2," etc.; (2) a dummy variable indicating whether the estimated conception date occurred before or after the earliest ZIP code spud date; (3) a continuous time variable indicating the number of months "with UNGD" (conception dates before the earliest spud date were assigned a value of 0, while conception dates after the earliest spud date were assigned a continuous time variable by month); and 4) a dummy variable assigned to each ZIP code indicating whether the earliest spud date occurred within the study period.

The following maternal characteristics were included in the final model: smoking status, education, race, age at delivery, pre-pregnancy body mass index, primary payer for delivery, WIC participation, diabetes, hypertension, and infection during pregnancy.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

Ma et al. 2016[1,2,3]

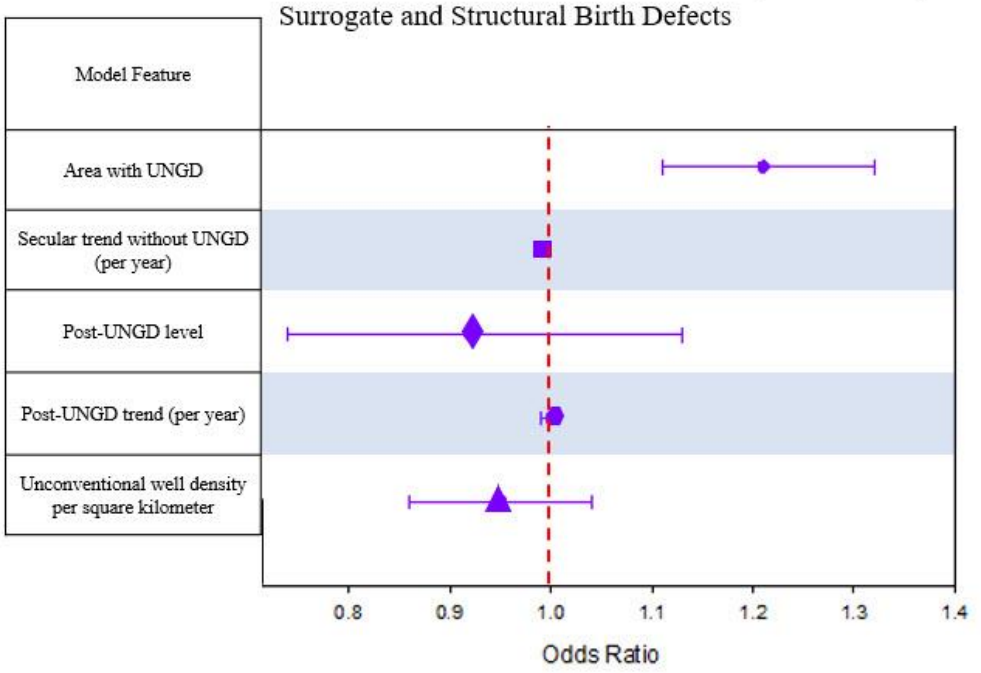
Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Any Birth Defect



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure surrogate: Total number of wells per kilometer squared calculated within respondent ZIP code. (Referent: Area without UNGD; n=1,424,170).

Ma et al. 2016 [1,2,3]

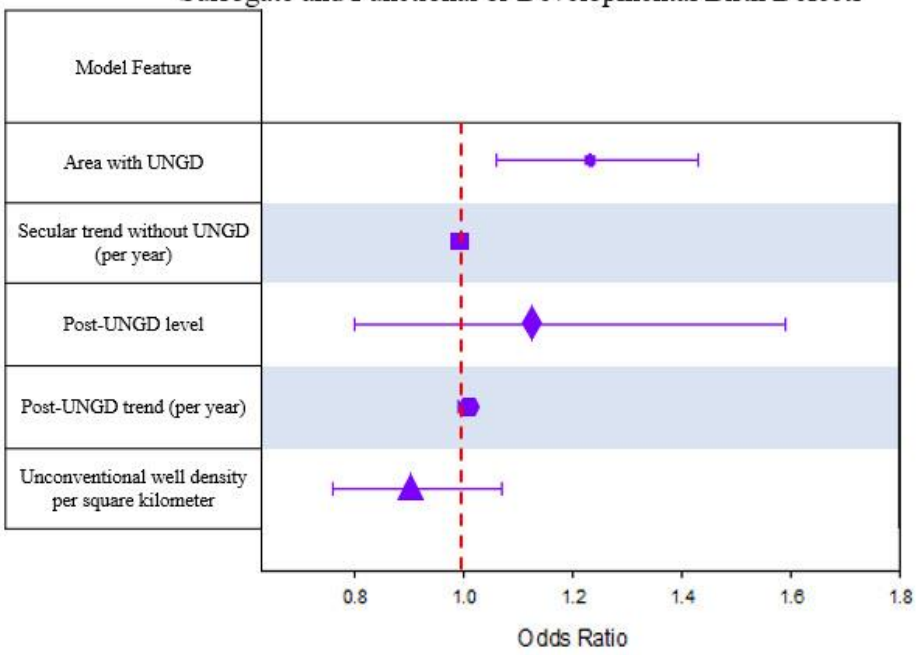
Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Structural Birth Defects



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure surrogate: Total number of wells per kilometer squared calculated within respondent ZIP code. (Referent: Area without UNGD; n=1,424,170).

Ma et al. 2016[1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Functional or Developmental Birth Defects



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure surrogate: Total number of wells per kilometer squared calculated within respondent ZIP code. (Referent: Area without UNGD; n=1,424,170).

Ma et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Study sample representative of the general Pennsylvania population.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on missing data, zip code or gestational age values.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to ecologic study design.	
	Control group appropriate to address study question	Not applicable to ecologic study design.	
	Same population over study period	Assessed this assumption in their analytical methods.	Unable to account for residential mobility.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: rates of hypertension.	Level of education, percent Black, percent private insurance, percent infection during pregnancy, age older than 35 higher in ZIP-codes categorized as without UOGD. Smoking during pregnancy, WIC use, diabetes and obesity prevalence lower in ZIP-codes categorized as without UOGD.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained outcome data from Pennsylvania birth certificate data.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	Potential reporting bias with respect to underreporting of birth defects in birth certificate data.
Exposure Assessment	Performed using valid, reliable and sensitive methods	Reliability depends on the quality of data available at PPADEP.	-Assumes residential stability throughout study period using address at delivery. -No discussion of quality of data.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Exposure assigned based on date of conception. -Included only unconventional wells.	- No evaluation of whether exposure surrogate represents UOGD activities. - No consideration of potential magnitude of exposure.
	Study period sufficient to capture exposure variability	Not applicable to ecologic study design.	

Ma et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Selection of exposure groups that represent the full range of variability in UOGD.		-Investigators do not test cut point bias. -Unexposed could be erroneously categorized as exposed if earliest spud date occurred after gestation but before birth.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included only wells post-spud date, and excluded wells with permits but not drilled.	Exposure metric may account for well that became inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Assigned during prenatal period using conception date.	
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data to collect ZIP-code level covariate information.	Control of confounding limited by available data on birth records.
	Controlled for baseline conditions	Controlled for basic measures of SES, demographics, and mother's BMI.	-No control of detailed SES or lifestyle factors. -Mother's diabetes status and infection during pregnancy collected but not controlled and has unbalanced proportions between exposure groups.
	Controlled for background exposures	Controlled for maternal smoking status before and during pregnancy.	No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells.
	Assessed time trends	-Control of time-trends using interrupted-time series methods. -Examined statistical interaction between year and areas with versus without UOGD.	
Analytical Methods	Analytical methods appropriate for study design	Yes. Multivariable segmented regression for interrupted time-series method.	Did not describe model-building procedure.
	Report measures of precision and variability	-Presented variability of maternal characteristics and outcome measures by exposure group. -95% Confidence intervals for odds ratios.	
	Report which statistical tests were used	-$p<0.05$ for significance testing.	

Ma et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Perform analysis to test sensitivity of results to alternative specifications.		No.
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Appropriate and complete interpretation of results	Investigators interpreted reported associations appropriately.	
	Discussion adequately addresses study limitations	-Discuss residential mobility, potential reporting bias. -Provide alternative explanations for findings	Do not discuss potential exposure misclassification or limitations of ecologic-level data.

Research Objective. The objective of this study was to examine the association between unconventional oil and gas well density at the county level and early infant mortality.

Study Period and Location. The study period was 1999 to 2014 in 67 Pennsylvania counties.

Study Population. The study population included all live births in Pennsylvania born between 1999 and 2014. Without defining the term "fracking," the investigators reported on the 10 counties with the highest number of "fracking" wells and split them into northwest (Susquehanna, Bradford, Wyoming, Lycoming, and Tioga) and a southwest (Washington, Westmoreland, Greene, Fayette, and Butler) groups. The investigators considered early infant mortality cases in other Pennsylvania counties as the referent.

Outcome Ascertainment. Live births and infant deaths data were obtained from the Pennsylvania Department of Health. Two outcomes were derived from these data: number of infant deaths between zero and 1 year and number of infant deaths between 0 and 28 days (early infant mortality).

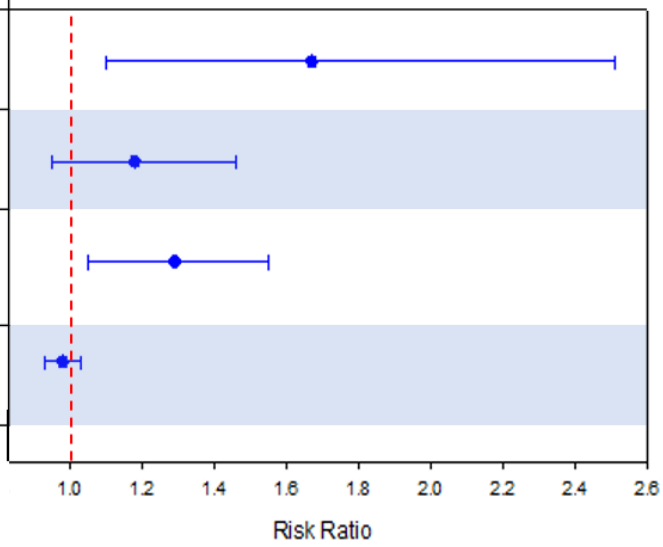
Exposure Assessment. "Fracking" well count and violations at the county level were obtained from the Pennsylvania Department of Environmental Protection. Non-fracking periods were considered to be 2003–2006 and the exposure period was considered to be 2007–2010. Additional analyses were conducted using exposure metrics of "water wells per birth" (obtained from the Pennsylvania department of Conservation and Natural Resources) and "violations per birth" in each county.

Analytical Methods and Covariate Inclusion. The investigators calculated unadjusted risk ratios of infant mortality in the first 28 days after birth by dividing the number of infant mortalities in 2007–2010 by the number of infant mortalities in 2003–2006 in each county. Separate analyses were performed calculating the risk ratio of infant mortality 0–28 days after birth for the full state, the full state less the ten "most fracked" counties, the northeastern most "fracked" counties, and the southwestern most "fracked" counties. The investigators also calculated the number water wells per unit birth. The investigators did not include any covariates in their models.

Results. The following figure summarizes results as presented by the study investigators (excluding any results provided in supplementary information).

Busby and Mangano et al. 2017 [1,2,3]
Risk Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Early Infant Mortality

Cases	Total *n*	Exposure Group
96	25,529	All Counties Northeast Pennsylvania
335	73,412	All Counties Southwest Pennsylvania
431	98,941	All 10 Northeast and Southwest PA counties
5,984	1,170,516	Entire State of Pennsylvania



1. Plotted from results presented in the study (excluding supplemental information).
2. Exposure surrogate: Total number of wells within the county of each study participant. (Referent: Non-fracked period [2003-2006]).
3. The data presented by the investigators in unadjusted. Adjusted data is not reported.

Busby and Mangano 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population		Unclear if investigators included all EIM records, or excluded records based on certain characteristics.
	Inclusion/Exclusion criteria specified		Not discussed by study investigators.
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to ecologic study design.	
	Control group appropriate to address study question	Not applicable to ecologic study design.	
	Same population over study period		-No formal assessment of this assumption over seven-year study period. - Unable to assess residential mobility because of study design.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)		Not assessed.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained outcome data from Pennsylvania vital birth and mortality statistics.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Reliability depends on the quality of data available at PADEP.	-Unclear which year investigators used to identify “fracked” counties. -No discussion of quality of data sources.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.		-No clear definition of "fracked" counties. -“Water wells per birth” analysis does not address question.
	Study period sufficient to capture exposure variability	Seven-year study period sufficient to capture variability over time.	
	Selection of exposure groups that represent the full range of variability in UOGD.		For temporal assessment: "unexposed" period (2003-2006) includes presence of UOGD in PA.

Busby and Mangano 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells		Not discussed by study investigators.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Yes (short at-risk period).	
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)		No potential countywide covariates assessed.
	Controlled for baseline conditions		No.
	Controlled for background exposures		No.
	Assessed time trends		No.
Analytical Methods	Analytical methods appropriate for study design		Analytical methods not discussed.
	Report measures of precision and variability		-Descriptive statistics not reported. -Confidence intervals selectively reported.
	Report which statistical tests were used		No.
	Perform analysis to test sensitivity of results to alternative specifications.		No.
Results and Discussion	All findings reported for analyses described in paper		Unclear.
	Appropriate and complete interpretation of results	Investigators suggest caution in interpreting analysis.	-Conclude that there is an effect, which is not consistent with reported results. -Reported risk ratio for all North East counties is reported in error.
	Discussion adequately addresses study limitations		No discussion of study limitations.

CURRIE ET AL. 2017

Research Objective. The objective of this study was to examine the association between residential proximity to UOGD development during gestation and perinatal health outcomes.

Study Period and Location. The study took place in Pennsylvania for the study period 2004–2013.

Study Population. The study population included all live singleton births in Pennsylvania between 2004 and 2013, excluding those missing outcome data, geocoded address at birth, maternal identifiers for record linkage, or duplicate records.

Outcome Ascertainment. The investigators obtained birth and early fetal health outcomes from Pennsylvania vital records.

Exposure Assessment. The investigators obtained data on all unconventional fractured wells from the Pennsylvania Department of Environmental Protection Internal Operator Well Inventory. They first calculated distance from the mother's residence at delivery to the nearest well. The "exposure" variable was the product of two variables: the "proximity" variable, which indicates whether the closest well from the maternal residence was within a specified distance, and the "timing" variable, which indicates whether the spud date of the closest well occurred before or after conception. In addition to the main "exposure" variable, the investigators also included a "near" variable, which indicated whether any wells were within a specified radius: 0–1, 1–2, 2–3, or 3–15 km.

Analytical Methods and Covariate Inclusion. Using *t*-tests, the investigators assessed whether population characteristics and outcome proportions differed before and after the first spud date for the population living 0–1 km from the nearest unconventional well and for the population living 3–15 km from the nearest well. The investigators then employed a difference-in-difference approach comparing covariates and outcomes before and after the first spud date in areas close (0–1 km) and far (3–15 km) from wells.

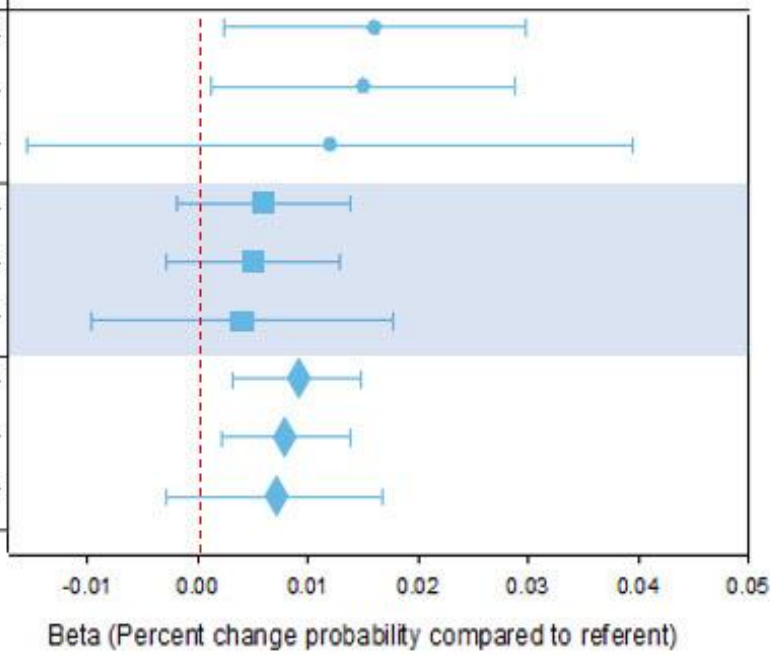
Next, the investigators created two regression models (Models 1 and 2), estimating probability of low birth weight and differences in birth weight or infant health index for mother *i* in year *t* (standard errors were clustered by mother). All models included the following variables: exposure, near (described above), county of mother's residence at delivery, delivery month and year, and linear time trend for one of six Pennsylvania regions. Both models also included several covariates. Model 1 included both time-varying and time-invariant covariates, including child gender, maternal race and ethnicity, mother's age, mother's education, marital status, and child parity. Model 2 included only time-varying covariates, such as mother's age, marital status, and child parity. Model 2 also included a variable representing a constant term for each mother's time-invariant covariates for mothers with multiple births during the study period. The investigators performed two iterations of each model: once with the full study population and again with only the study population living within 15 km of the nearest well.

Results. The following figure summarizes results as presented by the study investigators (excluding any results provided in supplementary information). The figures do not include infant health index results because the investigators did not provide these results numerically. For a comparison of continuous birth weight and preterm birth results across studies, see Figures 4-1 and 4-2, respectively in Section 4.1.2.

Currie et al. 2017 [1,2,3]

Betas and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Probability of Low Birth Weight

Model Feature	N	Radius Exposure Calculated
Full Population	1,117,919	2-3 km
	1,102,424	1-2 km
	1,086,917	0-1 km
Population <15 km	262,580	2-3km
	247,085	1-2km
	231,578	0-1km
Population <15km, fixed effect	262,580	2-3km
	247,085	1-2km
	231,578	0-1km



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Product of spud date and proximity (indicates if active well is within specified radius and indicates if the spud date of the closest well occurred before or after conception). Exposure Group: Continuous exposure.

Currie et al. 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Study sample representative of the general PA population.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on non-singleton births, missing data, and residential location.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to ecologic study design.	
	Control group appropriate to address study question	Not applicable to ecologic study design.	
	Same population over study period	Assessed this assumption analytically by examining mother's characteristics with multiple births over study period.	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: percent Hispanic, educational attainment, infant characteristics.	-Population living 0-1km from closest well: percent married and college degree higher “before” UOGD, and percent with age 20-24 higher “after” UOGD. -Population living 3-15 km from closest well: Percent black higher “before” UOGD.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained outcome data from Pennsylvania vital birth records.	Validity or clinical significance of "infant health index" unclear.
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Reliability depends on the quality of data available at PADEP.	-No discussion of geocoding methods. -Assumes residential stability throughout study period using address at delivery. -No discussion of quality of data source.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.

Currie et al. 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Assess exposure in a way that addresses review question.	Included only unconventional wells.	-No evaluation of whether exposure surrogate represents UOGD activities. - No consideration of potential magnitude of exposure.
	Study period sufficient to capture exposure variability	Not applicable to ecologic study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	Describe rationale for chosen radius to calculate exposure surrogate.	Investigators do not test cut-point bias.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included all active wells as of 2014.	May account for well that became inactive after 2014 or omit wells that become active after 2014.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Assigned during prenatal period based on date of conception.	Study includes period of no UOGD.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data to collect covariate information.	Control of confounding limited by available data on birth records.
	Controlled for baseline conditions	Control of basic demographic information.	-No control of co-morbidities or any SES factors. -No control of community-level factors.
	Controlled for background exposures		No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells.
	Assessed time trends	Model 2 includes mothers with multiple births during study period.	
Analytical Methods	Analytical methods appropriate for study design	- Linear regression (continuous outcomes), logistic regression (binary outcomes), accounting for individual- and community-level correlated standard errors. -Investigators describe model-building procedure.	
	Report measures of precision and variability	-Present variability of maternal characteristics and outcome measures over time. -Standard error reported for regression analyses.	
	Report which statistical tests were used	$p<0.01$, <0.05, <0.10 for significance testing.	

Currie et al. 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Perform analysis to test sensitivity of results to alternative specifications.	-Tested inclusion of inactive wells. -Tested exposure assignment based on birth date -Tested different distance specifications.	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Appropriate and complete interpretation of results	Investigators interpreted reported associations appropriately.	
	Discussion adequately addresses study limitations	-Described alternative explanations for findings. -Discussed some limitations of exposure assessment, small numbers in analytical models, potential benefits of UOGD, and limitations of available data.	Limited discussion of potential for residual confounding and bias.

WHITWORTH ET AL. 2017

Research Objective. The objective of this study was to examine the association between residential proximity to UOGD and perinatal outcomes.

Study Period and Location. This study took place in counties located in the Barnett Shale region in Texas, for the study period November 30, 2010 to November 29, 2012.

Study Population. The study population included 158,104 singleton births and 790 fetal deaths among women living in the study area. Investigators excluded births missing both the last menstrual period and clinical-based estimates of gestational age, births with estimated gestational age <22 or >44 weeks, births with implausible gestational age estimates, or fetal deaths with missing gestational age estimates. Births among women living greater than 20 miles from the nearest well were also excluded from the analysis.

Outcome Ascertainment. Outcomes of interest included: preterm birth, defined as a live birth delivered <37 weeks, small for gestation age (SGA), birth weight in grams, and fetal death. The investigators obtained these data from birth and fetal death records provided by the Texas Department of State Health Services.

Exposure Assessment. The investigators identified UNGD wells in the Barnett Shale with spud, completion, or production dates between January 1, 2010 and November 29, 2012. They included only unconventional wells and excluded wells with a permit date but no record of activity. The investigators then created an inverse distance weighted (IDW) exposure metric at three different radii around the home: 0.5, 2, and 10 miles. IDW values were then categorized into tertiles of exposure for each radius. Women with zero wells within 10 miles of their residence were considered the referent group.

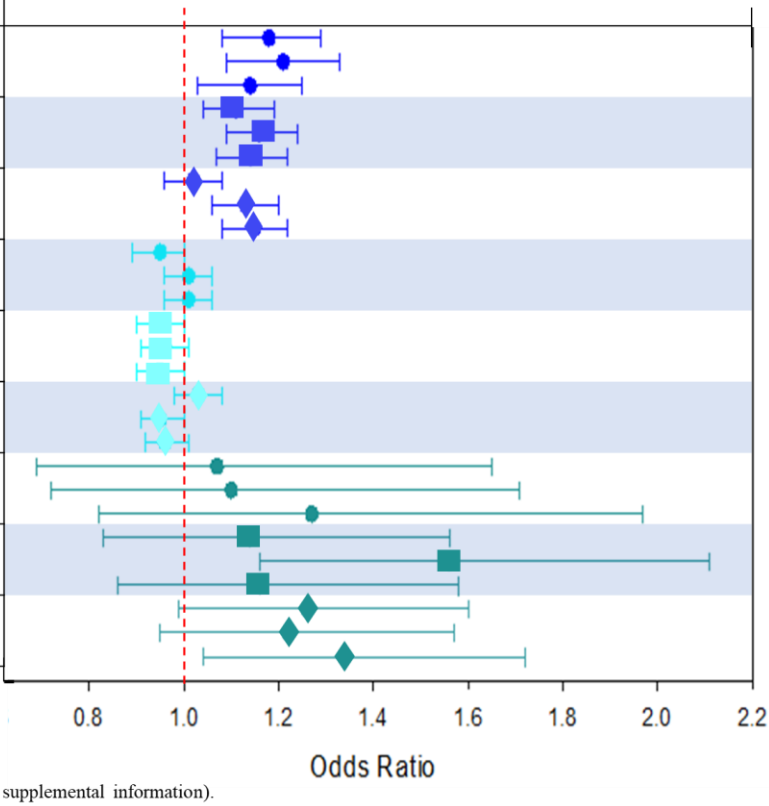
Analytical Methods and Covariate Inclusion. The investigators used multivariable logistic regression for all bivariate outcomes (preterm birth, SGA, and fetal death) and linear regression for birth weight. The investigators applied generalized estimating equations to all models, treating census tract as a random effect.

Results. The following figure summarizes results as presented by the study investigators (excluding any results provided in supplementary information). For a comparison of continuous birth weight and preterm birth results across studies, see Figures 4-1 and 4-2, respectively in Section 4.1.2.

Whitworth et al. 2017 [1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Adverse Perinatal Outcomes

End Point	Cases	Total *n*	Radius Exposure Calculated	Exposure Group
Preterm Birth	682	8,161	0.8 km	Tertile 1
	717	8,412		Tertile 2
	654	8,144		Tertile 3
	1,856	23,231	3.2 km	Tertile 1
	2,006	23,758		Tertile 2
	1,921	23,227		Tertile 3
	3,140	39,169	16.1 km	Tertile 1
	3,296	40,143		Tertile 2
	3,253	38,922		Tertile 3
SGA	964	8,161	0.8 km	Tertile 1
	1,063	8,409		Tertile 2
	1,013	8,142		Tertile 3
	2,785	23,227	3.2 km	Tertile 1
	2,984	23,757		Tertile 2
	2,847	23,223		Tertile 3
	5,233	39,168	16.1 km	Tertile 1
	4,924	40,139		Tertile 2
	4,877	38,917		Tertile 3
Fetal Death	27	8,188	0.8 km	Tertile 1
	29	8,438		Tertile 2
	32	8,174		Tertile 3
	74	23,301	3.2 km	Tertile 1
	103	23,860		Tertile 2
	77	23,300		Tertile 3
	157	39,325	16.1 km	Tertile 1
	138	40,277		Tertile 2
	149	39,066		Tertile 3



1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents a different outcome, and each shape represents a different model.
3. Exposure Surrogate: Inverse distance weighted-squared within 0.8, 3.2, and 16.1 kilometers of residence. Referent: Zero wells ≤ 10 miles from maternal residence.

Whitworth et al. 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Included all birth records from study area.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Excluded residence >20 miles from closest well, based on missing data, extreme or implausible gestational age estimates.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to retrospective cohort study design.	
	Control group appropriate to address study question	Not applicable to retrospective cohort study design.	
	Same population over study period	Brief study period (2010-2012)	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)		Population characteristics not reported by exposure group.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained using birth and fetal death records.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Reliability depends on the quality of underlying data (www.drillinginfo.com)	-Imprecise geocoding methods. -Assumes residential stability throughout study period using address at delivery. -No discussion of quality of data source.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Incorporated information about proximity and number of wells. -Exposure assigned over prenatal period. -Included only UNGD wells.	No evaluation of whether exposure surrogate represents UNGD activities.
	Study period sufficient to capture exposure variability	Not applicable to retrospective cohort study design.	

Whitworth et al. 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Selection of exposure groups that represent the full range of variability in UOGD.	Discussed rationale for distance chosen to calculate exposure.	Investigators do not test cut-point bias.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included only active UNGD wells in the Barnett Shale, and excluded UNGD wells that had not yet been drilled as reported in (drillinginfo.org).	May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Exposure assigned based on estimated conception date.	
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data to collect covariate information.	Control of confounding limited by available data on birth records.
	Controlled for baseline conditions	-Control of several prenatal variables: access to prenatal and medical care, pre-pregnancy BMI, pregnancy risk. -Controlled for basic demographic and SES characteristics.	-No control of co-morbidities or detailed SES factors. -No control of community-level factors.
	Controlled for background exposures	-Controlled for smoking status during pregnancy -Controlled for distance to major roadway in sensitivity analysis.	No control of other potential environmental sources, including occupational exposures, industrial sources, or conventional wells.
	Assessed time trends	Brief study period (2010-2012), so limited period for changing trends.	No assessment of time trends.
Analytical Methods	Analytical methods appropriate for study design	-Yes. Linear regression (continuous outcomes), logistic regression (binary outcomes). -Accounted for spatial correlation of women within Census tracts. -Investigators described model-building procedure.	
	Report measures of precision and variability	-Present variability of maternal characteristics and outcome measures by exposure group. -95% Confidence intervals for odds ratios and beta coefficients.	
	Report which statistical tests were used	$p<0.05$ for significance testing.	

Whitworth et al. 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Perform analysis to test sensitivity of results to alternative specifications.	-Tested residual confounding from traffic exposure, by controlling for distance to road in models. -Tested residual confounding from time-varying environmental factors by controlling for season of conception.	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Appropriate and complete interpretation of results	Investigators interpret reported findings accurately.	
	Discussion adequately addresses study limitations	Addresses limitations of exposure assignment and potential bias related to shorter exposure periods for SGA births and fetal deaths.	No discussion of potential for other background sources and other risk factors that may confound the association.

HILL 2018

Research Objective. The objective of this study was to assess the impact of shale gas development on infant health outcomes, primarily low birth weight (LBW), premature birth, and term birth weight (TBW).

Study Period and Location. This study took place in the Marcellus Shale area in Pennsylvania for the study period 2003 to 2010.

Study Population. The study population included singleton births in Pennsylvania from 2003 to 2010 (*n* = 1,098,884) for mothers with a residence within 2.5 km of a well.

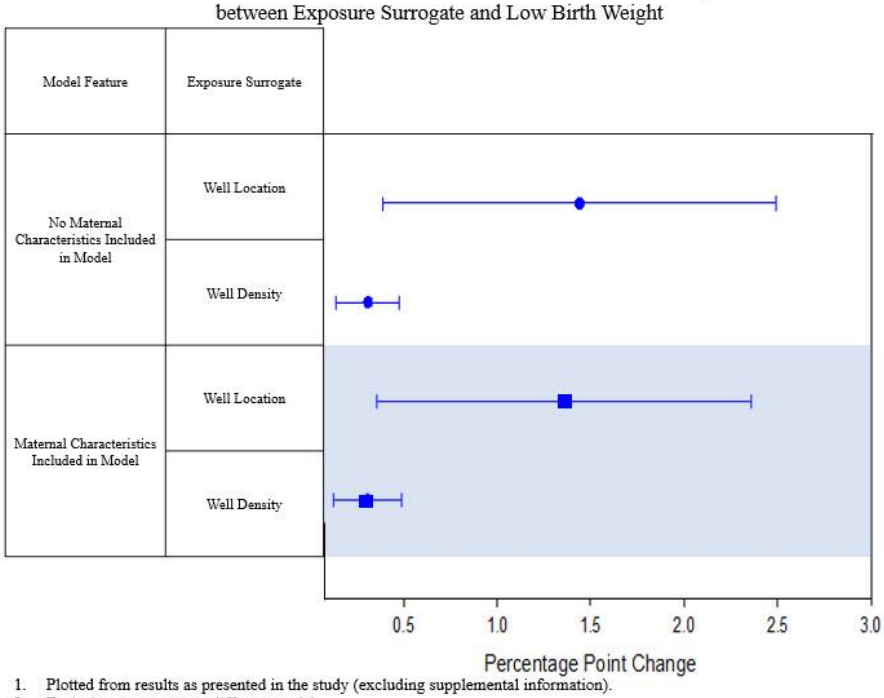
Outcome Ascertainment. Outcomes of interest included LBW, defined as birth weight less than 2500 g; premature birth, defined as gestation length less than 37 weeks; and TBW, defined as birth weight for infants who reach full term at 37 weeks (continuous measure). TBW is intended to study whether there was an average effect on the birth weight distribution as opposed to more extreme health outcomes (low birth weight and premature birth). The investigator also assessed alternative health measures for infant health, including birth weight (grams), APGAR score less than 8, gestation (weeks), small for gestational age, congenital anomaly, and a summary index (created to address the issue of precision). Investigator extracted natality and mortality data from restricted-access vital statistics for Pennsylvania. Extracted vital statistics included maternal characteristics (i.e., race, education, age, marital status, WIC status, insurance type, previous risky pregnancy, and smoking status).

Exposure Assessment. The investigator used a difference-in-differences model, in which mothers living within 2.5 km of a shale gas well before drilling commenced or permitted well served as controls for those women exposed after drilling began. Women with zero wells within 2.5 km of their residence were considered the referent group.

Analytical Methods and Covariate Inclusion. The investigator analyzed results using a difference-in-differences model. Investigator created a separate model for each outcome of interest. All models included indicators for month and year of birth, county of residence, drilling before birth (defined by closest well), any well or number of wells within 2.5 km of the residence, whether the birth occurred before the spud date of the closest well, and an interaction between the two latter variables. Maternal characteristics included in the analysis as abbreviated by the investigator were mother black, mother Hispanic, mother education, mother age, female child, WIC status, smoking during pregnancy, marital status, parity, previous risky pregnancy, and payment type. For each outcome, one regression was adjusted for maternal characteristics and one was not. Investigators examined the impact of well location and well density on birth outcomes and presented results as difference (percentage for binary outcomes and unit change for continuous outcomes) in the outcome compared to births with a spud date after birth.

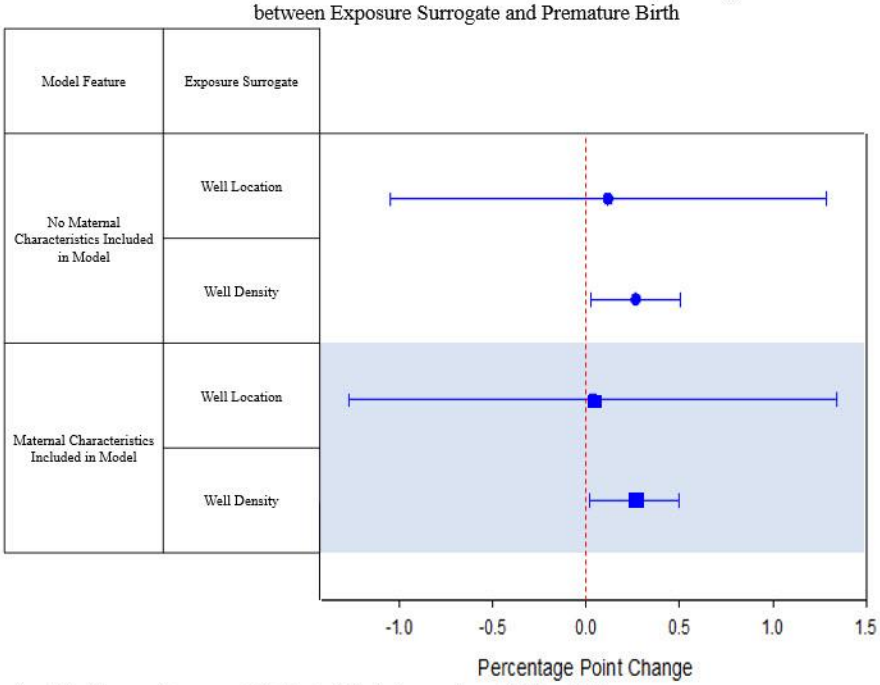
Results. For a comparison of continuous birth weight and preterm birth results across studies, see Figures 4-1 and 4-2, respectively in Section 4.1.2.

Hill 2018 [1,2,3]
Coefficients and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Low Birth Weight



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure surrogate: Product of two variables: 1) indicates if spud date of nearest well occurred before or after birth, and either 2a) indicates if active well is within specified radius (well location), or 2b) density of wells within specified radius (well density). N = 19,978.

Hill 2018[1,2,3]
Coefficients and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Premature Birth



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure surrogate: Product of two variables: 1) indicates if spud date of nearest well occurred before or after birth, and either 2a) indicates if active well is within specified radius (well location), or 2b) density of wells within specified radius (well density). N = 19,978.

Hill 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Representative of population living in areas near UOGD in Pennsylvania.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Excluded based on non-singleton births.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to retrospective cohort study design.	
	Control group appropriate to address study question	Not applicable to retrospective cohort study design.	
	Same population over study period	-Assessed maternal mobility for subsample of mothers with multiple births during the study period. -Assessed changes in maternal demographics over study period.	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: child sex, maternal education, maternal race and ethnicity, maternal marital status, smoking during pregnancy, and insurance type.	Lower proportion of older mothers and higher proportion of Medicaid and WIC recipients for births after versus before UOGD.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained from Pennsylvania vital birth and mortality statistics.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Reliability depends on the quality of data available at PADEP.	-Investigators do not discuss geocoding methods. -Assumes residential stability throughout prenatal period using address at delivery. -No discussion of quality of data source.
	Non-differential between outcome groups	Yes.	

Hill 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Incorporated information about proximity to wells. -Exposure assigned based on date of birth. -Included only "Marcellus shale wells."	No evaluation of whether exposure surrogate represents UOGD activities.
	Study period sufficient to capture exposure variability	Not applicable to retrospective cohort study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	Investigators describe motivation for chosen buffer distance (2.5 km) and test other distances.	
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included active Marcellus shale wells reported in PADEP database 2006-2010.	May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		No further consideration of timing of exposure with respect to gestational period, beyond birth date.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data to collect covariate information.	Control of confounding limited by available data on birth records.
	Controlled for baseline conditions	Controlled for basic SES, demographic factors, and history of risky pregnancy.	-No control of co-morbidities or detailed SES factors. -No control of community-level factors. -No control of prenatal care.
	Controlled for background exposures	Controlled for smoking.	No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells.
	Assessed time trends	-Detailed assessment of trends in select demographic characteristics over study period. -Controlled for month and year in models.	

Hill 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Analytical Methods	Analytical methods appropriate for study design	-Yes. -Takes into account correlations among siblings. -Investigators describe model-building procedure.	
	Report measures of precision and variability	-Present variability of maternal characteristics and outcome measures by exposure group. -Standard errors presented for coefficients, t-statistics for differences, and r-squared for model fit.	
	Report which statistical tests were used	-Difference-in-differences. -p<0.01, <0.05, <0.10 for significance testing.	No adjustment for multiple comparisons.
	Perform analysis to test sensitivity of results to alternative specifications.	-Tested well density as an exposure surrogate. -Tested various exposure surrogate distances. -Performed sensitivity analyses to test for residual confounding. -Performed subgroup analyses on the 10 most drilled and producing counties.	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Appropriate and complete interpretation of results	Investigators interpret reported main findings accurately.	No discussion of findings of lower risk of premature birth for mothers that live within 1 km vs. 2.5 km for closest well.
	Discussion adequately addresses study limitations		Weak discussion of study limitations.

JANITZ ET AL. 2018

Research Objective. The objective of this study was to evaluate whether living near natural gas wells is associated with critical congenital heart defects (CCHD), neural tube defects (NTD), and oral clefts in Oklahoma.

Study Period and Location. The study took place in Oklahoma for the study period January 1, 1997 to December 31, 2012.

Study Population. The study population included 476,600 singleton births in Oklahoma. The investigators evaluated all Oklahoma birth certificates between 1997 and 2009 and excluded births (1) that were outside of Oklahoma, in Osage County, non-geocodable, non-singleton, or having an address geocoded to ZIP code centroid or (2) that involved non-critical congenital heart defects, neural tube defects, and oral clefts. The analyses included 874 children with CCHDs, 217 children with NTDs, 603 children with oral clefts, and 474,935 children without anomalies.

Outcome Ascertainment. Outcomes of interest included CCHDs, NTDs (spina bifida and anencephaly), and oral clefts (cleft lip and cleft palate). All birth records without a linked congenital anomaly were classified as non-congenital anomalies. Investigators obtained data from the Oklahoma Birth Defects Registry (OBDR).

Exposure Assessment. The investigators obtained natural gas well data from the Oklahoma Corporation Commission for each year of study, including data on well location and monthly natural gas production levels. These data did not distinguish among different well types. Production data were used to determine whether a natural gas well was actively producing during each month and year of study. The investigators used an inverse distance-squared weighting (IDW) method to calculate the density of actively producing wells during the month of birth within a 2-, 5-, and 10-mile radius of the maternal residence at delivery. The primary analysis involved IDW-summed well counts divided into tertiles for the 2-mile radius, with sensitivity analyses conducted to evaluate the 5- and 10-mile radii. Mothers with no wells within a given radius were the referent group.

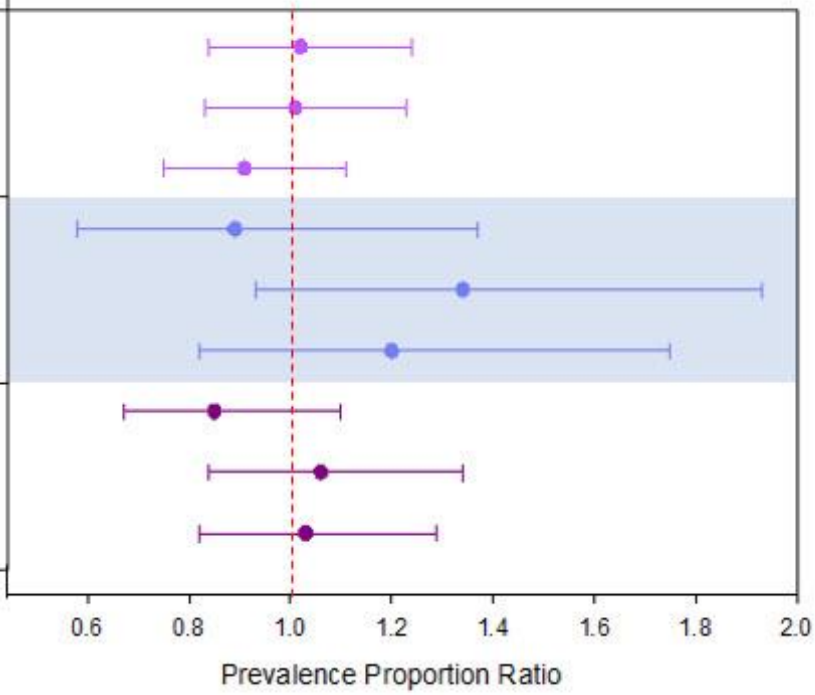
Analytical Methods and Covariate Inclusion. The investigators used modified Poisson regression with robust error variance to calculate prevalence proportion ratios, comparing children with and without congenital anomalies. The investigators used a directed acyclic graph (DAG) for model building. Based on the results, multivariable models were adjusted for maternal education only.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

Janitz et al. 2018 [1,2,3]

Prevalence Proportion Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and General Congenital Anomalies

End Point	Cases	Exposure Group
Critical Congenital Heart Defect	132	Tertile 1
	128	Tertile 2
	120	Tertile 3
Neural Tube Defects	28	Tertile 1
	39	Tertile 2
	36	Tertile 3
Oral Clefts	75	Tertile 1
	94	Tertile 2
	94	Tertile 3



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure Surrogate: Number of active natural gas wells with a 2 kilometer radius of maternal address; Referent: No wells within 2 kilometer radius of maternal address; 1,694 with anomalies, 474,935 without anomalies.

Janitz et al. 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Included all live singleton births in all "birthing hospitals" in Oklahoma.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Excluded residence outside of study area, addresses not geocoded, non-singleton births, and birth defects that were not the focus of study.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to retrospective cohort study design.	
	Control group appropriate to address study question	Not applicable to retrospective cohort study design.	
	Same population over study period		No formal assessment of this assumption over 12-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: race/ethnicity, maternal age, parity, smoking during pregnancy, urbanization.	Cases were higher proportion male (CCHD and oral cleft only), more frequently born at an earlier gestational age and lower birthweight.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained birth defects data from the OBDR	Potential reporting bias with respect to underreporting of birth defects in birth certificate data (OBDR includes diagnoses between 2 and 6 years of age).
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	-Reliability depends on the quality of data available at Oklahoma Corporation Commission. -Precise geocoding methods.	-Assumes residential stability throughout study period using address at delivery. -No discussion of quality of data source.

Janitz et al. 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Non-differential between outcome groups	Well information collected independently of outcome data	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	Incorporated information about proximity to and number of wells (objective to assess exposure to all natural gas wells in study area).	-No evaluation of whether exposure surrogate represents OGD activities. -Did not differentiate between unconventional and conventional wells. -Included period before rapid onset of UOGD.
	Study period sufficient to capture exposure variability	Not applicable to retrospective cohort study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	Evaluated various buffer distances (2-, 5-, and 10-mile).	-Did not test cut-point bias. -No justification for choice of radius within which to calculate exposure surrogate.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included all active natural gas wells (defined as having production data at least one month per year).	May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Tested exposure surrogate assigned to the month of conception.	Exposure assigned using month of birth (rather than day).
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	-Used electronic health record data to collect covariate information on basic child characteristics, maternal factors, prenatal care -Categorized residences as urban or rural using Census data.	
	Controlled for baseline conditions	-Controlled only for maternal education in primary models. -Fully adjusted models of anomaly categories included child and mother demographic factors, prenatal care, tobacco use, and education.	-No control of co-morbidities or detailed SES factors. -No control of community-level factors. -No fully adjusted model assessed for specific anomalies.
	Controlled for background exposures		No control of potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells.

Janitz et al. 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Assessed time trends		No.
Analytical Methods	Analytical methods appropriate for study design	-Yes: multivariable Poisson regression with robust error variance. -Described model-building procedures. -Used DAG software to determine inclusion of covariates in model.	
	Report measures of precision and variability	-Present variability of maternal characteristics and outcome measures by exposure group. -Measures of uncertainty with confidence intervals around effect estimates.	
	Report which statistical tests were used	Prevalence proportion ratios and 95% confidence intervals	
	Perform analysis to test sensitivity of results to alternative specifications.	-Tested a fully adjusted model. -Tested various exposure specifications (e.g., different buffer distances, IDW-squared). -Tested using date of conception (rather than birth date) for exposure assignment. -Tested excluding addresses geocoded to the street.	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Appropriate and complete interpretation of results	Investigators interpret reported findings accurately.	
	Discussion adequately addresses study limitations	Nuanced discussion of limitations in exposure and outcome data, residential mobility, potential for selection bias, small number of cases, and limited covariate data.	-Do not describe alternative explanations for findings. - No discussion of multiple comparisons.

WHITWORTH ET AL. 2018

Research Objective. The objective of this study was to examine the association between exposure to UNGD and preterm birth and to explore effects by trimester of exposure, preterm birth severity, and UNGD phases.

Study Period and Location. This study took place in counties located in the Barnett Shale regions of Texas, for the study period November 30, 2010 to November 29, 2012.

Study Population. The study population included 166,966 singleton births among women living in the study area. Investigators excluded births missing both the last menstrual period and clinical-based estimates of gestational age, births with estimated gestational age <22 or >44 weeks, births with implausible birth weight for gestational age estimates or births whose maternal residence geocoded location mapped outside of the study area.

Outcome Ascertainment. Outcomes of interest included preterm birth, defined as a live birth delivered <37 weeks. Investigators also stratified preterm birth by World Health Organization clinical definitions of preterm-birth severity. The investigators obtained these data from birth records provided by the Texas Department of State Health Services.

Exposure Assessment. The investigators used a multistage process to characterize exposure to wells. First, the investigators identified wells in the study area with spud, completion, or production dates between January 1, 2010 and November 29, 2012. The investigators included only UNGD wells (as opposed to conventional wells) and excluded wells with a permit date but no record of activity. The investigators then created inverse distance weighted (IDW) exposure metrics for wells located within 0.5 mile of the mother's residence at delivery. The IDW metric was calculated separately for drilling and production phases and was divided into tertiles over the full pregnancy and separately for each trimester. Women with zero wells within 0.5 mile of their residence were considered the referent.

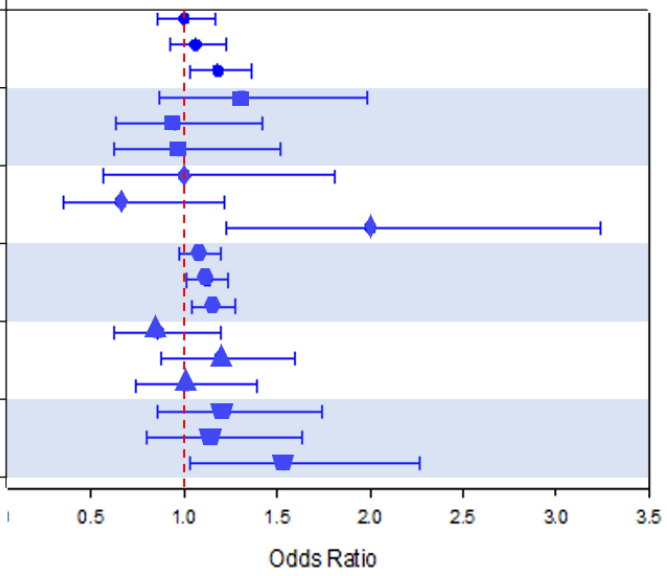
Analytical Methods and Covariate Inclusion. Cases (n = 13,332) were matched to controls (66,933) by maternal age at delivery and race/ethnicity. The investigators used multivariable logistic regression for the primary association of interest (preterm birth), and polytomous regression for the preterm birth severity analysis. The following covariates were collected for model inclusion: maternal education, parity, smoking during pregnancy, pre-pregnancy body mass index, infant sex, previous poor pregnancy outcome, timing and frequency of prenatal care, and maternal residential distance to the nearest major roadway. Pre-pregnancy body mass index, education, smoking, infant sex, previous poor pregnancy outcome, and prenatal care utilization index were included in final models.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information). For a comparison of preterm birth results across studies, see Figures 4-1 and 4-2, respectively in Section 4.1.2.

Whitworth et al. 2018 [1,2,3]

Odds Ratios with 95% Confidence Intervals for the Relationship between Exposure Surrogate and Preterm Birth by Preterm Birth Severity

Phase	Preterm Birth Severity	Cases	Total *n*	Exposure Group
Drilling	Moderately Preterm	237	1,813	Tertile 1
		250	1,831	Tertile 2
		287	1,912	Tertile 3
	Very Preterm	31	1,813	Tertile 1
		32	1,831	Tertile 2
		27	1,912	Tertile 3
	Extremely Preterm	15	1,813	Tertile 1
		13	1,831	Tertile 2
		28	1,912	Tertile 3
Production	Moderately Preterm	480	3,502	Tertile 1
		501	3,519	Tertile 2
		543	3,621	Tertile 3
	Very Preterm	48	3,502	Tertile 1
		61	3,519	Tertile 2
		54	3,621	Tertile 3
	Extremely Preterm	49	3,502	Tertile 1
		46	3,519	Tertile 2
		38	3,621	Tertile 3

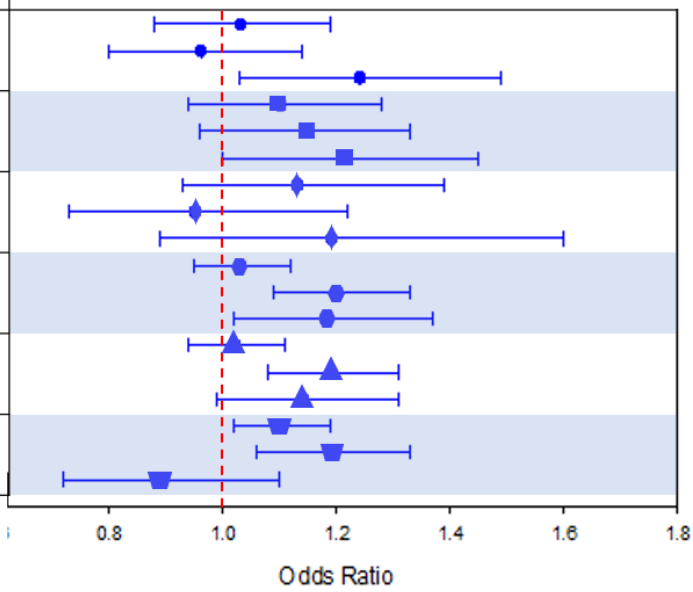


1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Inverse distance weighted-squared within 0.8 kilometers of residence.

Whitworth et al. 2018 [1,2,3]

Odds Ratios with 95% Confidence Intervals for the Relationship between Exposure Surrogate and Preterm Birth by Trimester of Exposure

Phase	Model Feature	Cases	Total *n*	Exposure Group
Drilling	Trimester 1	227	1,482	Tertile 1
		169	1,116	Tertile 2
		167	898	Tertile 3
	Trimester 2	233	1,380	Tertile 1
		183	1,077	Tertile 2
		161	878	Tertile 3
	Trimester 3	134	799	Tertile 1
		82	527	Tertile 2
		63	358	Tertile 3
Production	Trimester 1	791	4,937	Tertile 1
		558	3,081	Tertile 2
		252	1,408	Tertile 3
	Trimester 2	820	5,135	Tertile 1
		594	3,324	Tertile 2
		279	1,574	Tertile 3
	Trimester 3	1,024	5,994	Tertile 1
		451	2,508	Tertile 2
		109	745	Tertile 3



1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Inverse distance weighted-squared within 0.8 kilometers of residence.

Whitworth et al. 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Included all birth records from study area.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on missing data, implausible data, or address not geocoded.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to case-control study design.	
	Control group appropriate to address study question	-Clearly delineated between cases (preterm birth) and controls (not preterm birth). -Selected controls from study population.	
	Same population over study period	Brief study period (2010-2012).	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: SES, lifestyle factors, and demographic factors.	Adequate prenatal care utilization more prevalent among cases than controls.
Outcome Assessment	Outcome ascertained using valid and reliable measures	-Ascertained using birth records and employed two different estimates of gestational age. -Adjusted time-at-risk for controls based on matched cases. -Corrected implausible birthweight for gestational age values.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Dependent on the quality of underlying data (www.drillinginfo.com).	-Street-level geocoding. -Assumed residential stability throughout study period. -No discussion of data quality.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.

Whitworth et al. 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Assess exposure in a way that addresses review question.	-Incorporated information about proximity to and number of wells. -Exposure assigned over prenatal period. -Included only UNGD wells. -Assigned at daily resolution.	No evaluation of whether exposure surrogate represents UOGD activities.
	Study period sufficient to capture exposure variability	Not applicable to case-control study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	Discussed rationale for distance chosen to calculate exposure.	Did not test cut-point bias.
	Differentiates among UOGD and its various phases	Reflected drilling and production phases.	Did not reflect well pad construction and well completion phases.
	Differentiates between active and non-active wells	Included only active UNGD wells in the Barnett Shale.	
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Exposure assignment overlapped prenatal period, and with separate assignments by trimesters.	
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	-Used electronic health record data to collect covariate information. -Matched by maternal age at delivery and race/ethnicity.	Control of confounding limited by available data on birth records.
	Controlled for baseline conditions	Detailed control of prenatal risk factors.	-No control of detailed lifestyle or detailed SES factors. -No assessment of potential confounding by community-level factors.
	Controlled for background exposures	Controlled for distance to major roadway.	No control of other potential environmental sources, including occupational exposures, industrial sources, or conventional wells.
	Assessed time trends	Brief study period (2010-2012), so limited period for changing trends.	
Analytical Methods	Analytical methods appropriate for study design	-Yes. -Described model-building procedure.	
	Report measures of precision and variability	-Presented variability of maternal characteristics by case status. -95% confidence intervals for odds ratios.	
	Report which statistical tests were used	-Odds ratios for conditional logistic regression.	

Whitworth et al. 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
		-$p<0.05$ for model building; $p<0.01$ for trend across tertiles.	
	Perform analysis to test sensitivity of results to alternative specifications.	Polytomous regression for preterm severity.	No other sensitivity analyses.
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Appropriate and complete interpretation of results	Interpreted reported effect estimates appropriately.	
	Discussion adequately addresses study limitations	Detailed discussion of limitations around phase differentiation, sources of potential selection bias, residential mobility and potential exposure routes.	No discussion of other potential environmental sources or explanations for observed associations.

MOKRY 2010

Research Objective. The objective of this study was to examine cancer incidence in two Flower Mound, Texas ZIP codes in response to citizen concerns about gas drilling in their communities.

Study Period and Location. The study took place in Flower Mound, Texas. Cancer incidence was compared between two time periods: 1998–2007 and 2007–2009.

Study Population. The study population included residents of the two ZIP codes included in the study area.

Outcome Ascertainment. The investigators obtained childhood leukemia, non-Hodgkin's lymphoma, breast cancer, and childhood brain/central nervous system (CNS) case data for the years 1998–2009 from the Department of State Health Services Texas Cancer Registry. The Census 2000 population data were used as the comparison population in calculating the standardized incidence ratio (SIR).

Exposure Assessment. The investigators compared SIRs between two time periods: 1998–2007 and 2007–2009.

Analytical Methods and Covariate Inclusion. The investigators calculated age-standardized, sex-stratified SIRs, adjusting for race, and 99% confidence intervals of childhood leukemia, non-Hodgkin's lymphoma, breast cancer, and childhood CNS cancers for the two ZIP codes included in the study. Numbers of cancer cases were compared between two time periods: 1998–2007 and 2007–2009.

Results. Investigators did not present study results separately for the two periods. Please see the report for numeric SIRs and confidence intervals, presented by age, sex, and ZIP code for all cancers assessed for the 1998–2007 time period.

Mokry 2010: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Cases representative of study population of included ZIP-codes.	
	Inclusion/Exclusion criteria specified	Restricted analysis to population with specific cancers and residence of 2 ZIP-codes.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to ecologic study design.	
	Control group appropriate to address study question		Used Statewide data for "expected" cases of ZIP-code cancer cases.
	Same population over study period	Discussed population growth over study period.	No formal assessment of this assumption over 10-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)		Not assessed.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained by identifying tumors in statewide cancer registry.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Yes.	
Exposure Assessment	Performed using valid, reliable and sensitive methods		-Assumed residential stability throughout study period. -County categorizations and choice of temporal periods not clear.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	Partially addressed study objective to assess whether cancer incidence was elevated before UOGD using temporal assessment.	-Study design did not allow investigators to distinguish among well types. -No consideration of potential magnitude of exposure.
	Study period sufficient to capture exposure variability		Study did not include period of high UOGD activity.
	Selection of exposure groups that represent the full range of variability in UOGD.		Low UOGD during "exposed" period.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells		No.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		Timeframe not sufficient given expected latency for cancers and study period assessed.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure		None collected.

Mokry 2010: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	groups (cohort studies) or cases and controls (case-control studies)		
	Controlled for baseline conditions	Expected cancer rates provided by age, sex, and race.	No control of baseline conditions beyond basic demographics.
	Controlled for background exposures		No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells.
	Assessed time trends		Not assessed.
Analytical Methods	Analytical methods appropriate for study design	Yes.	Qualitative comparisons between "exposed" and "unexposed" periods.
	Report measures of precision and variability	99% Confidence intervals for SIRs.	Variability in characteristics over time not presented.
	Report which statistical tests were used	Standardized incidence ratios.	No statistical assessment of difference between time periods.
	Perform analysis to test sensitivity of results to alternative specifications.		No.
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Appropriate and complete interpretation of results	-Interpreted reported effect estimates appropriately given findings and study limitations. -Discussed other explanations for findings.	
	Discussion adequately addresses study limitations	Mentioned small sample size and population mobility.	-No discussion of limitations of ecologic-level assessment. -No discussion of consideration of latency or potential for exposure misclassification. -No discussion of lack of control of confounding.

FRYZEK ET AL. 2013

Research Objective. The objective of this study was to compare childhood cancer incidence before and after the first well was drilled in each county in Pennsylvania.

Study Period and Location. The study location consisted of the entire Commonwealth of Pennsylvania between 1990 and 2009.

Study Population. The study population consisted of children younger than 20 years of age who had been diagnosed with any form of cancer between 1990 and 2009.

Outcome Ascertainment. Investigators obtained childhood cancer cases from 1990–2009 from the Pennsylvania Cancer Registry. Expected rates of cancer in the general population were provided by the Surveillance, Epidemiology, and End Results (SEER) program. Outcomes of interest included all childhood cancer cases, childhood leukemia cases, and central nervous system (CNS) tumors.

Exposure Assessment. Publicly accessible spud data reports for all well types from 1990–2009 were obtained from the Pennsylvania Department of Environmental Protection. The first spud date in each county was used to estimate "pre" (1990–first spud date in county) versus "post" (first spud date–2009 in county) oil and gas development exposure. Exposure was also categorized using county well density as no wells (referent), 1–500 wells, 501–1000 wells, 1001–2000 wells, and >2001 wells.

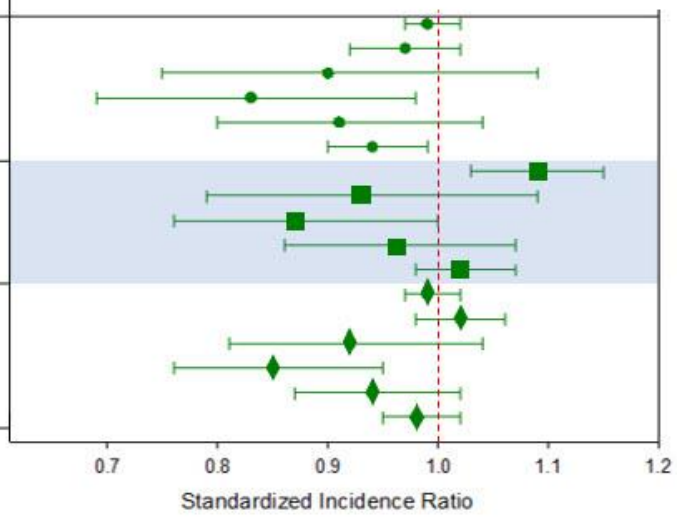
Analytical Methods and Covariate Inclusion. The investigators calculated standardized incidence ratios (SIRs) using summed observed and expected childhood cancers, leukemia, and CNS tumors for each county by sex, race, and age group. Statewide SIRs were calculated separately for "pre" and "post" drilling periods. SIRs were also stratified by number of wells and type of wells (gas, oil, horizontal gas wells, and Marcellus shale wells).

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

Fryzek et al. 2013 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and All Cancers

Time Period	Observed Cases	Exposure Group
Before Drilling	6,838	No wells
	1,403	1-500 wells
	114	501-1,000 wells
	129	1,001-2,000 wells
	228	>2,001 wells
	1,874	Total for all counties w/wells
After Drilling	1,291	1-500 wells
	153	501-1,000 wells
	202	1,001-2,000 wells
	350	>2,001 wells
	1,996	Total for all counties w/wells
Total Study Period (1990-2009)	6,838	No wells
	2,694	1-500 wells
	267	501-1,000 wells
	331	1,001-2,000 wells
	578	>2,001 wells
	3,870	Total for all counties w/wells

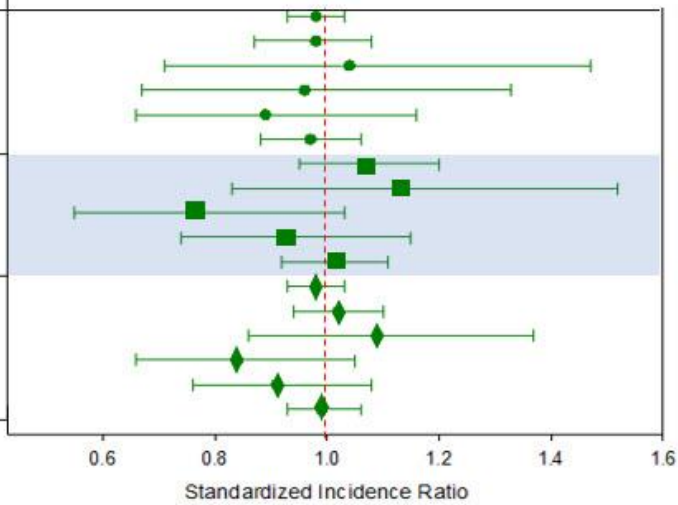


1. Plotted from results as presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure Surrogate: Outcome rate before or after earliest spud date within county of residence. (Referent: No wells in county; n = Population of Pennsylvania).

Fryzek et al. 2013 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Leukemia

Time Period	Observed Cases	Exposure Group
Before Drilling	1,640	No wells
	339	1-500 wells
	31	501-1,000 wells
	35	1,001-2,000 wells
	52	>2,001 wells
	457	Total for all counties w/wells
After Drilling	303	1-500 wells
	45	501-1,000 wells
	42	1,001-2,000 wells
	81	>2,001 wells
	471	Total for all counties w/wells
Total Study Period (1990-2009)	1,640	No wells
	642	1-500 wells
	76	501-1,000 wells
	77	1,001-2,000 wells
	133	>2,001 wells
	928	Total for all counties w/wells

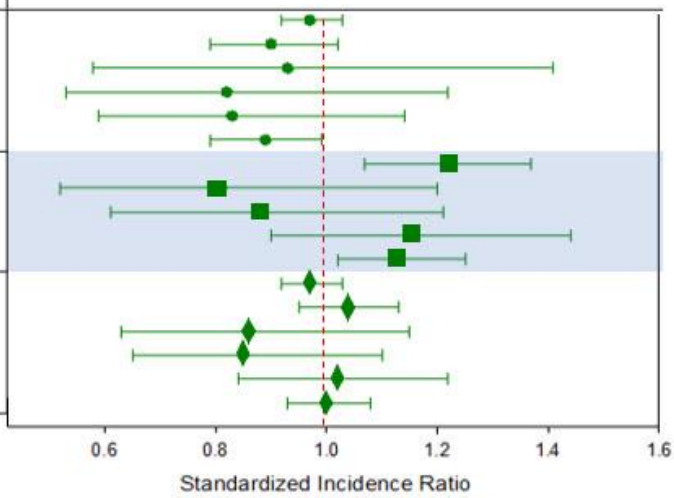


1. Plotted from results as presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure Surrogate: Outcome rate before or after earliest spud date within county of residence. (Referent: No wells in county).

Fryzek et al. 2013 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Central Nervous System Tumors

Time Period	Observed Cases	Exposure Group
Before Drilling	1,225	No wells
	242	1-500 wells
	22	501-1,000 wells
	24	1,001-2,000 wells
	39	>2,001 wells
	327	Total for all counties w/wells
After Drilling	257	1-500 wells
	24	501-1,000 wells
	36	1,001-2,000 wells
	75	>2,001 wells
	392	Total for all counties w/wells
Total Study Period (1990-2009)	1,225	No wells
	499	1-500 wells
	46	501-1,000 wells
	60	1,001-2,000 wells
	114	>2,001 wells
	719	Total for all counties w/wells

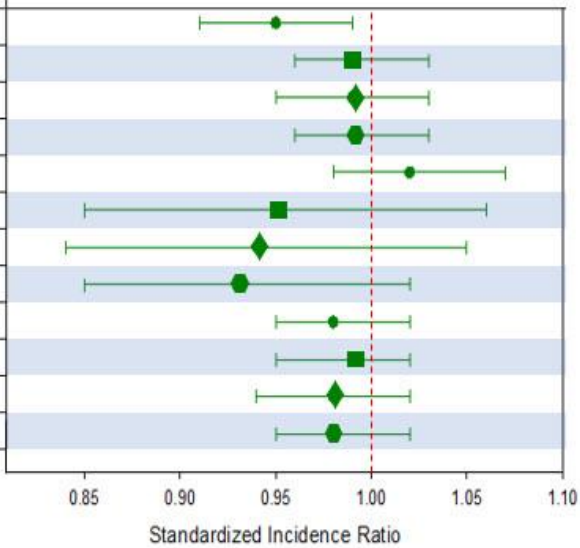


1. Plotted from results as presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure Surrogate: Outcome rate before or after earliest spud date within county of residence. (Referent: No wells in county).

Fryzek et al. 2013 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and All Cancers, Stratified by Well Type

Time Period	Well Type	Observed Cases
Before Drilling	Gas Wells	1,942
	Horizontal Wells	2,754
	Horizontal Gas Wells	2,429
	Marcellus Shale Wells	2,542
After Drilling	Gas Wells	1,928
	Horizontal Wells	341
	Horizontal Gas Wells	313
	Marcellus Shale Wells	506
Total Study Period (1990-2009)	Gas Wells	3,870
	Horizontal Wells	3,095
	Horizontal Gas Wells	2,742
	Marcellus Shale Wells	3,048

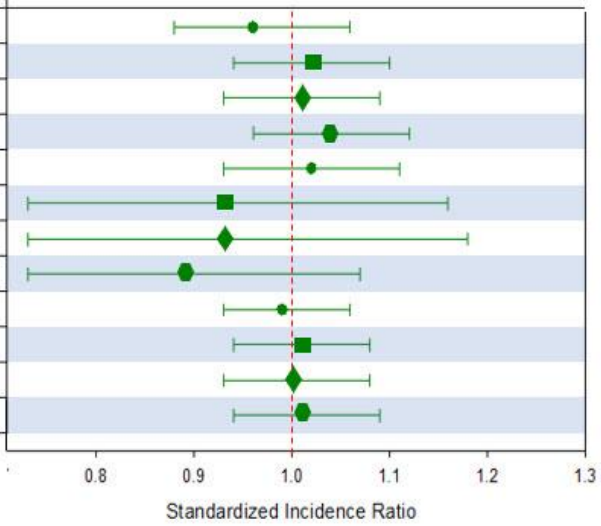


1. Plotted from results as presented in the study (excluding supplemental information).
2. Each color represents a different outcome, and each shape represents a different model.
3. Exposure Surrogate: Outcome rate before or after earliest spud date within county of residence. (Referent: No wells in county).

Fryzek et al. 2013[1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Leukemia, Stratified by Well Type

End Point	Model Feature	Observed Cases
Before Drilling	Gas Wells	470
	Horizontal Wells	679
	Horizontal Gas Wells	596
	Marcellus Shale Wells	635
After Drilling	Gas Wells	458
	Horizontal Wells	76
	Horizontal Gas Wells	71
	Marcellus Shale Wells	112
Total Study Period (1990-2009)	Gas Wells	928
	Horizontal Wells	755
	Horizontal Gas Wells	667
	Marcellus Shale Wells	747

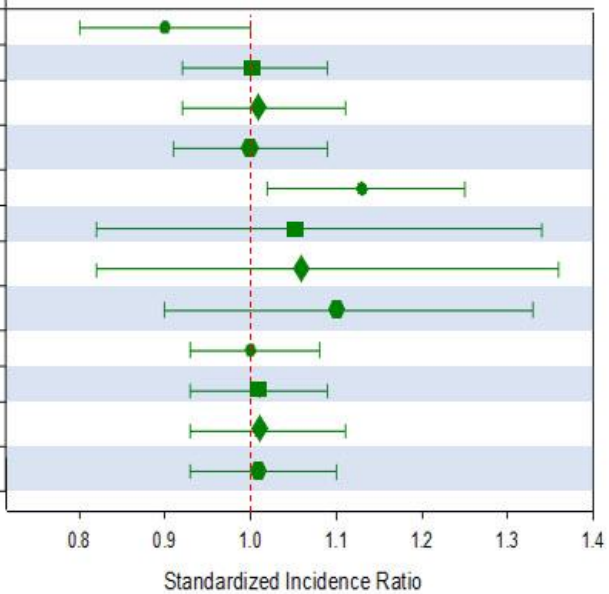


1. Plotted from results as presented in the study (excluding supplemental information).
2. Each color represents a different outcome, and each shape represents a different model.
3. Exposure Surrogate: Outcome rate before or after earliest spud date within county of residence. Comparison group: before spud date. Time period: before versus after spud date.

Fryzek et al. 2013 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Central Nervous System Tumors, Stratified by Well Type

End Point	Model Feature	Observed Cases
Before Drilling	Gas Wells	340
	Horizontal Wells	509
	Horizontal Gas Wells	454
	Marcellus Shale Wells	468
After Drilling	Gas Wells	379
	Horizontal Wells	66
	Horizontal Gas Wells	62
	Marcellus Shale Wells	104
Total Study Period (1990-2009)	Gas Wells	719
	Horizontal Wells	575
	Horizontal Gas Wells	516
	Marcellus Shale Wells	572



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each color represents a different outcome, and each shape represents a different model.
3. Exposure Surrogate: Outcome rate before or after earliest spud date within county of residence. Comparison group: before spud date.

Fryzek et al. 2013: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Cases representative of general population.	
	Inclusion/Exclusion criteria specified	Exclusions based on age.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to ecologic study design.	
	Control group appropriate to address study question	Used the general population in each county to calculate expected cancer cases.	
	Same population over study period		-No assessment of this assumption over 20-year study period. -Applied 1990 and 2000 Census population estimates to estimate population sizes across study period (1990-2009).
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)		Not reported.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained by identifying tumors in statewide cancer registry using ICD-9 codes.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Outcome recorded without knowledge of exposure.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Dependent on quality of underlying data (PADEP)	-Assumed residential stability throughout study period. -No discussion of quality of PADEP data.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.		-Differentiated by well type in a secondary analysis. -Objective to examine cancer rates in relation to UOGD activities, but only 2.5% of wells in study area were unconventional wells.
	Study period sufficient to capture exposure variability		Study period includes few potential years at risk to UOGD.
	Selection of exposure groups that represent the full range of variability in UOGD.		Did not test for cut point bias for stratified analysis.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included only wells post-spud date.	May account for wells that become inactive later in study period.

Fryzek et al. 2013: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		Timeframe not sufficient given expected latency for cancers and study period assessed.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)		None collected.
	Controlled for baseline conditions	Expected cancer rates by age, sex, and race.	No control of baseline conditions beyond basic demographics.
	Controlled for background exposures		No control of other potential environmental sources, including industrial sources, traffic, or conventional wells.
	Assessed time trends		Not assessed.
Analytical Methods	Analytical methods appropriate for study design	Yes. Standardized incidence ratios (SIRs).	Qualitative comparisons between "exposed" and "unexposed" periods.
	Report measures of precision and variability	95% Confidence intervals for SIRs.	Variability in characteristics over time not presented.
	Report which statistical tests were used	No statistical assessment of difference between time periods.	
	Perform analysis to test sensitivity of results to alternative specifications.		No.
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Discussion adequately addresses study limitations	Mentioned small sample size and potential for residual confounding.	-No discussion of limitations of ecologic-level assessment. -No discussion of consideration of latency or potential for exposure misclassification. -No discussion of no control of potential confounding.
	Appropriate and complete interpretation of results	Interpreted reported effect estimates appropriately.	Investigator conclusions do not reflect results, given limitations.

FINKEL 2016

Research Objective. The objective of this study was to examine the association between unconventional oil and gas development and cancer incidence at the county level.

Study Period and Location. The study took place in six southwestern Pennsylvania counties (Allegheny, Beaver, Fayette, Greene, Washington, and Westmoreland) and township-level data for Washington County. Investigators collected cancer and well activity data for the years 2000–2012.

Study Population. The study population included residents of the six counties included in the study area.

Outcome Ascertainment. The investigators obtained urinary bladder, thyroid, and leukemia cancer cases (defined as the number of primary tumors, not individuals) and cancer deaths by age, sex, race, and primary site from the Pennsylvania Cancer Registry. The Pennsylvania Cancer Registry also provided county-level standardized incidence ratios.

Exposure Assessment. The investigators assumed an exposure period from 2008–2012. They classified the six counties as high, medium, or low with respect to the number of oil-producing wells. The cut points for these classifications were not defined in the study.

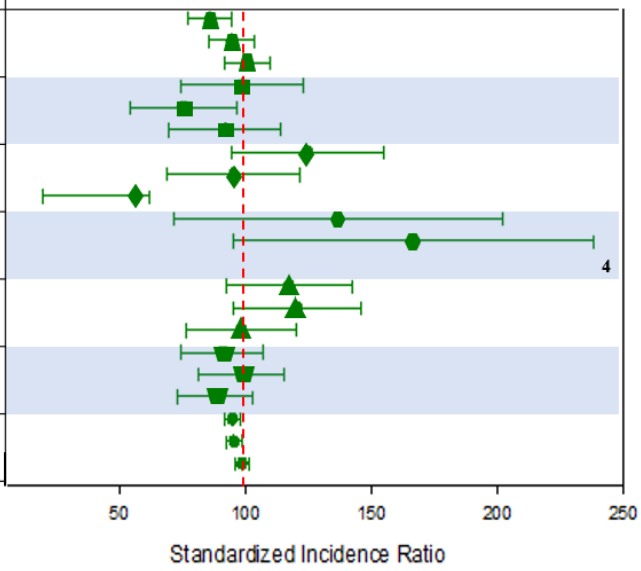
Analytical Methods and Covariate Inclusion. The investigators calculated age-standardized incidence ratios by sex and 95% confidence intervals of urinary bladder cancer, thyroid cancer and leukemia for each of the six counties for three different time periods: 2000–2004, 2004–2008, and 2008–2012. The investigators also calculated percent change in cancer incidence between the 2000–2004 and 2008–2012 periods.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

Finkel et al. 2016 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Leukemia in Females

County	Observed Cases	Exposure Group
Allegheny County Only	388	2000-2004
	425	2004-2008
	478	2008-2012
Beaver County Only	63	2000-2004
	48	2004-2008
	64	2008-2012
Fayette County Only	66	2000-2004
	51	2004-2008
	31	2008-2012
Greene County Only	17	2000-2004
	21	2004-2008
	7	2008-2012
Washington County Only	84	2000-2004
	88	2004-2008
	79	2008-2012
Westmoreland County Only	120	2000-2004
	132	2004-2008
	130	2008-2012
All PA	3,741	2000-2004
	3,874	2004-2008
	4,426	2008-2012

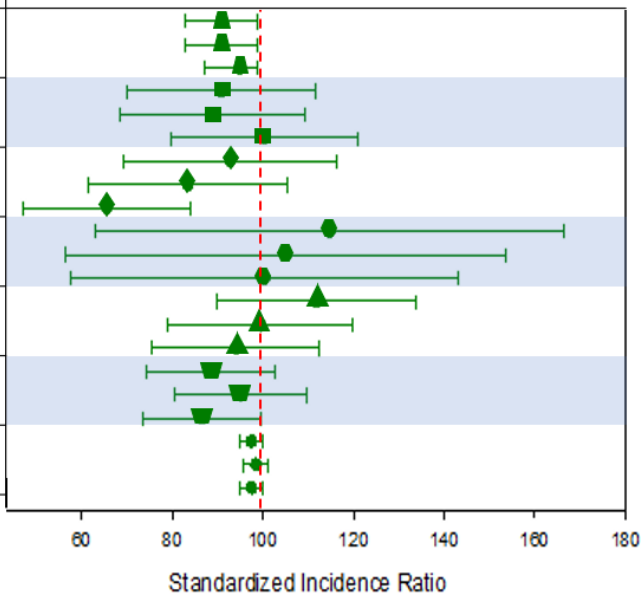


1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Comparison between time periods.
4. No data provided in study.

Finkel et al. 2016 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Leukemia in Males

County	Observed Cases	Exposure Group
Allegheny County Only	490	2000-2004
	487	2004-2008
	559	2008-2012
Beaver County Only	74	2000-2004
	73	2004-2008
	91	2008-2012
Fayette County Only	60	2000-2004
	55	2004-2008
	47	2008-2012
Greene County Only	19	2000-2004
	18	2004-2008
	21	2008-2012
Washington County Only	101	2000-2004
	92	2004-2008
	100	2008-2012
Westmoreland County Only	149	2000-2004
	163	2004-2008
	170	2008-2012
All PA	4,833	2000-2004
	5,049	2004-2008
	5,747	2008-2012

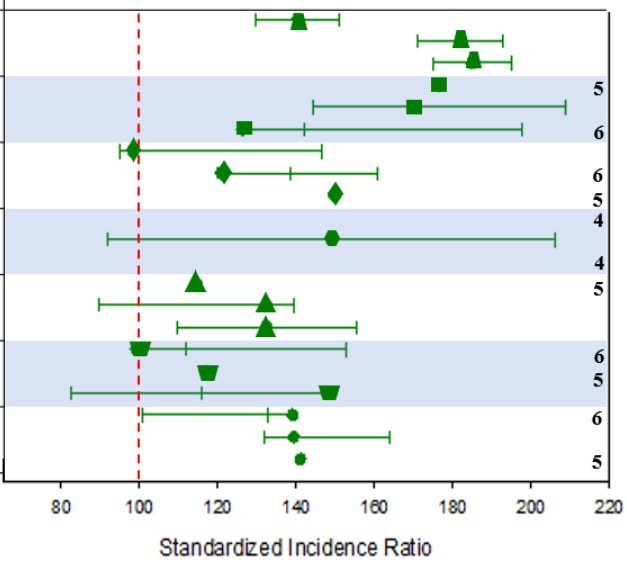


1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Comparison between time periods.

Finkel et al. 2016 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Thyroid Cancer in Females

County	Observed Cases	Exposure Group
Allegheny County Only	646	2000-2004
	1,074	2004-2008
	1,345	2008-2012
Beaver County Only	115	2000-2004
	145	2004-2008
	131	2008-2012
Fayette County Only	52	2000-2004
	84	2004-2008
	124	2008-2012
Greene County Only	9	2000-2004
	26	2004-2008
	4	2008-2012
Washington County Only	84	2000-2004
	129	2004-2008
	163	2008-2012
Westmoreland County Only	133	2000-2004
	206	2004-2008
	325	2008-2012
All PA	5,969	2000-2004
	7,950	2004-2008
	10,237	2008-2012

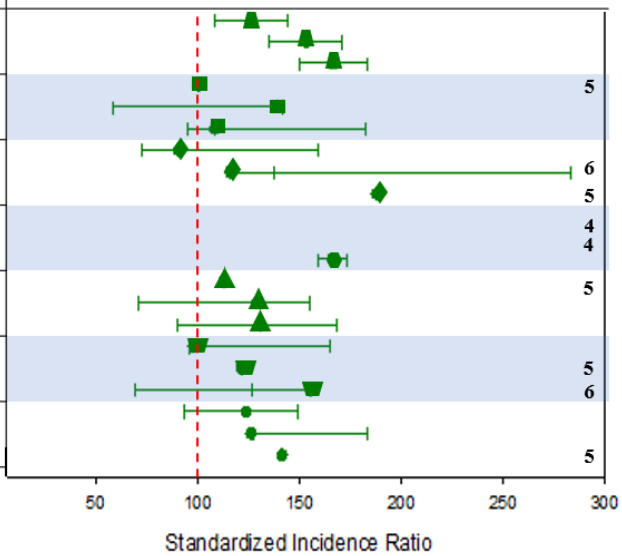


1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Comparison between time periods.
4. No data provided in study.
5. No 95% confidence intervals provided.
6. Results as presented appear to be in error.

Finkel et al. 2016 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Thyroid Cancer in Males

County	Observed Cases	Exposure Group
Allegheny County Only	186	2000-2004
	284	2004-2008
	383	2008-2012
Beaver County Only	22	2000-2004
	39	2004-2008
	37	2008-2012
Fayette County Only	16	2000-2004
	27	2004-2008
	53	2008-2012
Greene County Only	2	2000-2004
	6	2004-2008
	13	2008-2012
Washington County Only	28	2000-2004
	42	2004-2008
	54	2008-2012
Westmoreland County Only	45	2000-2004
	73	2004-2008
	117	2008-2012
All PA	1,752	2000-2004
	2,341	2004-2008
	3,350	2008-2012

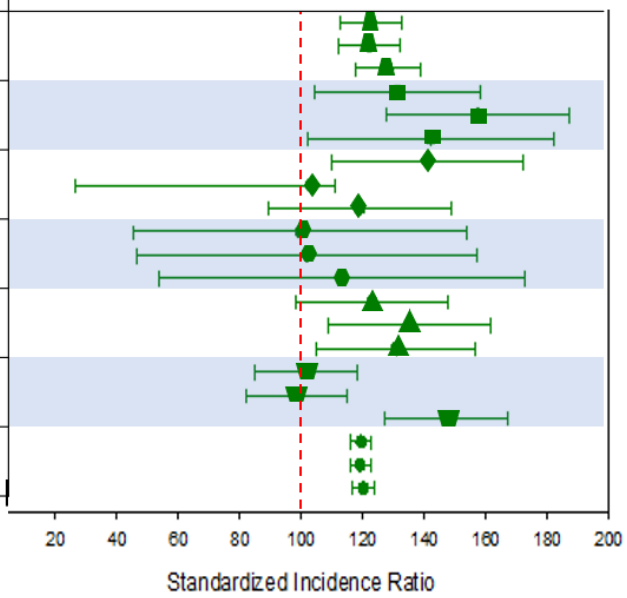


1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Comparison between time periods.
4. No data provided in study.
5. No 95% confidence intervals provided.
6. Results as presented appear to be in error.

Finkel et al. 2016 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Urinary Bladder Cancer in Females

County	Observed Cases	Exposure Group
Allegheny County Only	589	2000-2004
	561	2004-2008
	570	2008-2012
Beaver County Only	90	2000-2004
	106	2004-2008
	94	2008-2012
Fayette County Only	80	2000-2004
	57	2004-2008
	62	2008-2012
Greene County Only	13	2000-2004
	13	2004-2008
	14	2008-2012
Washington County Only	94	2000-2004
	101	2004-2008
	99	2008-2012
Westmoreland County Only	144	2000-2004
	137	2004-2008
	208	2008-2012
All PA	4,917	2000-2004
	4,831	2004-2008
	4,941	2008-2012

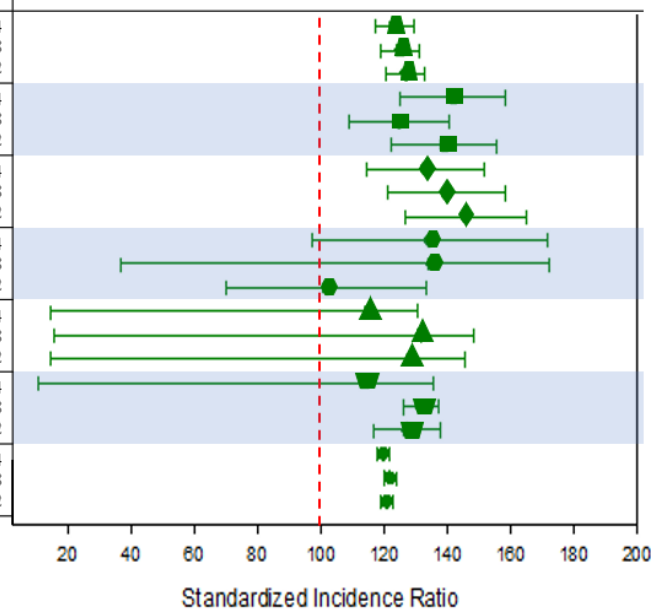


1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Comparison between time periods.

Finkel et al. 2016 [1,2,3]

Standardized Incidence Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Urinary Bladder Cancer in Males

County	Observed Cases	Exposure Group
Allegheny County Only	1,547	2000-2004
	1,555	2004-2008
	1,588	2008-2012
Beaver County Only	273	2000-2004
	242	2004-2008
	273	2008-2012
Fayette County Only	201	2000-2004
	214	2004-2008
	223	2008-2012
Greene County Only	50	2000-2004
	52	2004-2008
	40	2008-2012
Washington County Only	243	2000-2004
	283	2004-2008
	292	2008-2012
Westmoreland County Only	452	2000-2004
	532	2004-2008
	542	2008-2012
All PA	13,496	2000-2004
	14,149	2004-2008
	14,843	2008-2012



1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Comparison between time periods.

Finkel et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Cases representative of population from study area.	
	Inclusion/Exclusion criteria specified	Included cancer cases from specified counties.	Inclusion and exclusion criteria not specified.
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to ecologic study design.	
	Control group appropriate to address study question	Used the general population in each county to calculate expected cancer cases.	
	Same population over study period		No assessment of this assumption over 12-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar among included counties: age distribution.	Percent change in population from 2010 to 2014, percent White, and percent poverty differed among counties
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained by identifying tumors in statewide cancer registry using ICD-0 codes.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Yes.	
Exposure Assessment	Performed using valid, reliable and sensitive methods		-Assumed residential stability throughout study period. -County categorizations and choice of temporal periods not clear.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	Assessed whether cancer incidence was elevated before UOGD using temporal assessment.	-Study design did not allow for investigators to distinguish well type. - No consideration of potential magnitude of exposure.
	Study period sufficient to capture exposure variability	12-year study period sufficient to capture variability over time and overlaps with period of high UOGD.	
	Selection of exposure groups that represent the full range of variability in UOGD.		-Number of wells used to categorize counties as "high," "medium," and "low" not reported -Unclear why other counties were excluded form analysis.

Finkel et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
			-UOGD occurred during "unexposed" period.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Presented number of active unconventional wells in counties.	Investigators did not quantify exposure based on wells.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Timeframe sufficient for cancers with short latencies.	Timeframe not sufficient for cancers with long latencies.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)		None collected.
	Controlled for baseline conditions	Expected cancer rates provided by age, sex, and race.	No control of baseline conditions beyond basic demographics.
	Controlled for background exposures		No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells.
	Assessed time trends		Not assessed.
Analytical Methods	Analytical methods appropriate for study design	Yes. Standardized incidence ratios.	Qualitative comparisons between "exposed" and "unexposed" periods.
	Report measures of precision and variability	95% confidence intervals for standardized incidence ratios.	-Variability in characteristics over time not presented. -Reported 95% confidence intervals appear in be an error
	Report which statistical tests were used		No discussion of tests used to assess significant differences between time periods.
	Perform analysis to test sensitivity of results to alternative specifications.		No.
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Discussion adequately addresses study limitations	Discussed sample size, population change, potential confounding, and limitations of ecologic-level assessment.	
	Appropriate and complete interpretation of results	Appropriate interpretation of reported associations.	Provided contradictory conclusions based on interpretations of results and study limitations.

MCKENZIE ET AL. 2017

Research Objective. The objective of this paper was to determine if residential proximity to oil and gas wells was associated with increased risk of childhood hematologic cancer.

Study Period and Location. The study took place in rural areas and towns in Colorado with less than 50,000 residents. The study period was between 1991 and 2013, with health data collected from 2001–2013 and well data collected for 1991–2013.

Study Population. The study population consisted of children age 0–24 years of age in the Colorado Central Cancer Registry who resided in the study area at the time of diagnosis between 2001 and 2013. Cases were excluded if they lacked a geocoded address. Cases (n = 215) were defined as children diagnosed with acute lymphocytic leukemia (ALL) or non-Hodgkin's lymphoma (NHL). Controls were children diagnosed with other cancers (n = 528) in the study area.

Outcome Ascertainment. Cancer incidence data were obtained from the Colorado Department of Public Health and Environment's Central Cancer Registry. The registry also provided data on child demographic characteristics.

Exposure Assessment. The investigators identified "active" wells (between spud-in and abandon dates) through the Colorado Oil and Gas Information System. Investigators used the earliest recorded well activity date as surrogates for missing spud dates. Distance between wells and residential addresses were calculated within 16.1 km of the residence to produce an IDW "active" well count metric. IDW well count was averaged over 5 years in the 1–5 years prior to diagnosis for subjects 0–4 years of age, 10 years in the 1–10 years prior to diagnosis for subjects 5–19 years of age, and 5 years in the 6–10 years prior to diagnosis for subjects 20–24 years of age. The study population was divided into tertiles of IDW well count. Those living greater than 16.1 km from an active well were considered the referent group.

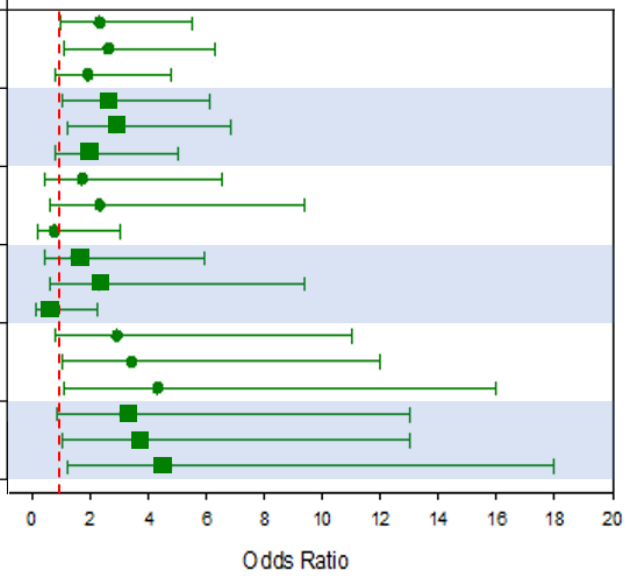
Analytical Methods and Covariate Inclusion. The investigators used multivariable logistic regression to examine associations between IDW well count tertiles and odds of ALL or NHL cancers. The following covariates were included in adjusted models: child age group (5-year intervals), gender, race/ethnicity, ZIP-code income quintile, and residential elevation (Model 1). A second model (Model 2) included year of cancer diagnosis in addition to Model 1 covariates. The investigators performed secondary analyses examining IDW well count within an 8-km (rather than 16.1-km) radius around each residence and adjusting for smoking in the subset of participants with maternal smoking data.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

McKenzie et al. 2017[1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Acute Lymphocytic Leukemia

Population Assessed	Cases (% of total study pop.)	Controls (N)	Model Feature	Exposure Group
Total Study Population (Ages 0-24)	21 (14%)	132	Adjusted for age, race, gender, SES, and elevation.	Tertile 1
	26 (18%)	119		Tertile 2
	25 (16%)	130		Tertile 3
	21 (14%)	132	Adjusted for age, race, gender, SES, elevation, and year of diagnosis.	Tertile 1
	26 (18%)	119		Tertile 2
	25 (16%)	130		Tertile 3
Ages 0-4	12 (33%)	24	Adjusted for race, gender, SES, and elevation.	Tertile 1
	11 (35%)	20		Tertile 2
	9 (21%)	34		Tertile 3
	12 (33%)	24	Adjusted for race, gender, SES, elevation and year of diagnosis.	Tertile 1
	11 (35%)	20		Tertile 2
	9 (21%)	34		Tertile 3
Ages 5-24	9 (7.7%)	108	Adjusted for age, race, gender, SES, and elevation.	Tertile 1
	15 (14%)	99		Tertile 2
	16 (14%)	96		Tertile 3
	9 (7.7%)	108	Adjusted for age, race, gender, SES, elevation, and year of diagnosis.	Tertile 1
	15 (14%)	99		Tertile 2
	16 (14%)	96		Tertile 3

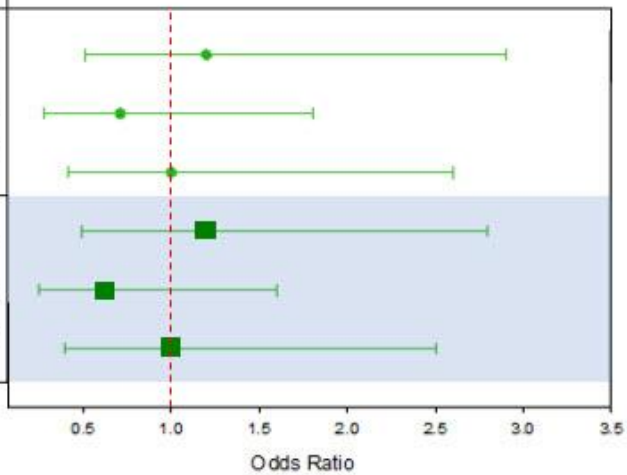


1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Inverse distance weighted within 16.1 kilometers of residence. Referent: No wells within 16.1 kilometers of residence.

McKenzie et al. 2017[1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Non-Hodgkins Lymphoma

Model Feature	Cases	Exposure Group
Adjusted for age, race, gender, SES, and elevation.	132	Tertile 1
	119	Tertile 2
	130	Tertile 3
Adjusted for age, race, gender, SES, elevation, and year of diagnosis.	132	Tertile 1
	119	Tertile 2
	130	Tertile 3



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Inverse distance weighted within 16.1 kilometers of residence (Controls n = 147). Referent: No wells within 16.1 kilometers of residence.

McKenzie et al. 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Included all cancer hematologic cancer cases in the study area.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on age, cancer type, and residence. Residences that could not be geocoded	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to case-control study design.	
	Control group appropriate to address study question	Control group derived from the same study population as controls.	Control group may have different risk factors than cases.
	Same population over study period		No formal assessment of this assumption over 12-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: percent Non-Hispanic White, higher elevation, ZIP-code income.	Percent non-White, male, age 5-14 higher in Non-Hodgkin's Lymphoma cases than controls. Percent non-white, male, low elevation, younger age groups higher in Acute Lymphocytic Lymphoma cases than controls.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Derived from state cancer registry.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Yes.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	–Dependent on quality of underlying data Colorado Oil and Gas Information System (COGIS). -Precise geocoding.	-No discussion of quality of COGIS. -No control of mobility during exposure period.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	Incorporated information about proximity to and number of wells (Objective to assess exposure to all natural gas wells in study area).	-Did not evaluate whether exposure surrogate represents OGD activities. -Included period before rapid onset of UOGD.

McKenzie et al. 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
			-Did not distinguish between unconventional and conventional wells.
	Study period sufficient to capture exposure variability	Not applicable to case-control study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.		-Did not test for cut point bias. -No justification for choice of radius within which to calculate exposure surrogate.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included all active oil and gas wells reported in Colorado Oil and Gas Information System (COGIS).	May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Included latency periods and lag times based on age of diagnosis.	
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used cancer registry to collect covariate information.	
	Controlled for baseline conditions	Controlled for basic demographic, SES factors, and year of diagnosis.	-No control of baseline conditions beyond basic demographics. -No control of community-level factors.
	Controlled for background exposures	-Restricted to rural population to reduce potential impact of urban exposures. -Controlled for residential elevation.	No control of other potential environmental sources, including industrial sources, traffic, or conventional wells.
	Assessed time trends		No.
Analytical Methods	Analytical methods appropriate for study design	Yes. Logistic regression.	
	Report measures of precision and variability	-Presented variability of maternal characteristics by exposure and outcome groups -95% Confidence intervals for odds ratios.	
	Report which statistical tests were used	Yes.	

McKenzie et al. 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Perform analysis to test sensitivity of results to alternative specifications.	-Tested sensitivity of alternative buffer distances around maternal residence. -Tested impact of maternal smoking	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Discussion adequately addresses study limitations	-Mentioned small sample size, potential for residual confounding, residential mobility, lack of hydrologic and meteorological information, missing geocoded addresses.	No discussion of choice of buffer distance.
	Appropriate and complete interpretation of results	Interpreted reported effect estimates appropriately.	

RASMUSSEN ET AL. 2016

Research Objective. The objective of this paper was to evaluate associations between UOGD and asthma exacerbation.

Study Period and Location. The study area included portions of New York and Pennsylvania in the catchment area for the Geisinger Clinic. The study period was 2005 to 2012.

Study Population. The study population included all patients ages 5 to 90 with asthma who had contact with the Geisinger Clinic during the study period. Patients with cystic fibrosis, chronic pulmonary heart disease, paralysis of vocal cords or larynx, bronchiectasis, and pneumoconiosis were excluded. Investigators matched cases to controls by age, sex, and year of encounter.

Outcome Ascertainment. Asthma exacerbations were defined as a new oral corticosteroid medication order (mild exacerbation), asthma emergency department visit (moderate exacerbation), and asthma hospitalizations (severe exacerbation). The investigators identified controls as patients diagnosed with asthma within Geisinger Health System on any day within the study period. Under this approach, cases could become controls.

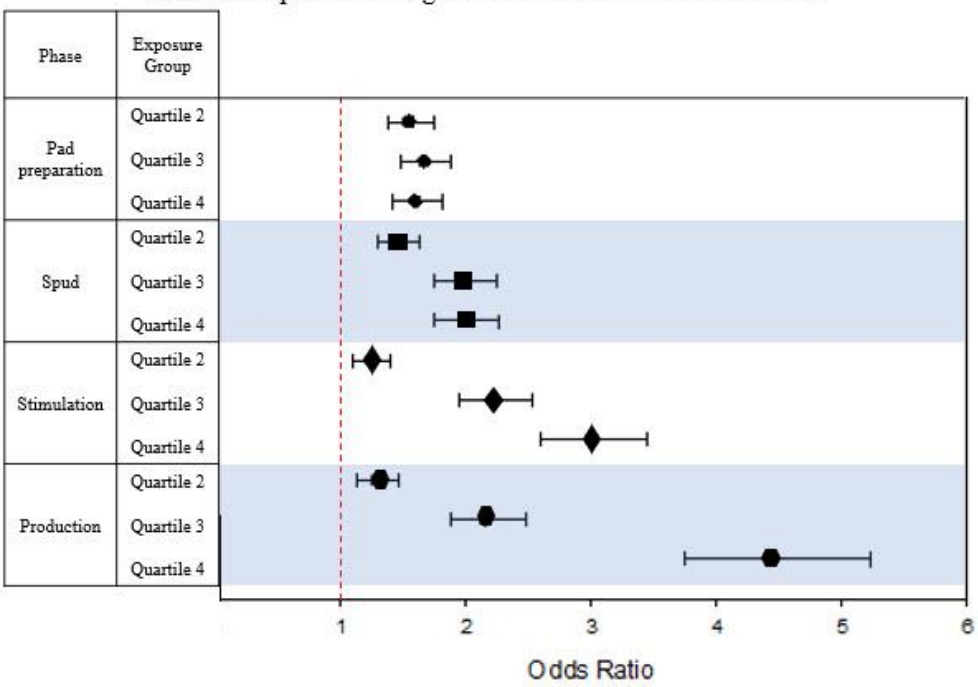
Exposure Assessment. Well spud, production, stimulation dates, and drilling depth data from 2005–2012 were collected from the Pennsylvania Department of Environmental Protection and the Pennsylvania Department of Conservation and Natural Resources. Crowd-sourced photos from SkyTruth were used to determine well location. The investigators developed an inverse-distance weighted (IDW) exposure surrogate for four UOGD phases. The surrogate incorporated distance between wells and residence, well depth, and production volume in the day before contact with the Geisinger Clinic (index date). The four exposure metrics were then categorized into quartiles (very low activity [referent], low activity, medium activity, and high activity).

Analytical Methods and Covariate Inclusion. The investigators used multilevel logistic regression with random intercept for patient and community to account for multiple events within each patient and spatial correlation within communities. Time-varying covariates included in the model were age at event, season of the event, smoking status, overweight status, Medical Assistance (i.e., Medicaid), Type 2 diabetes for each index date, and distance to nearest major and minor road using road networks maintained by the Federal Highway Administration. Non-time varying covariates included were sex, race/ethnicity, family history of asthma, and community-defined socioeconomic deprivation.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

Rasmussen et al. 2016 [1,2,3]

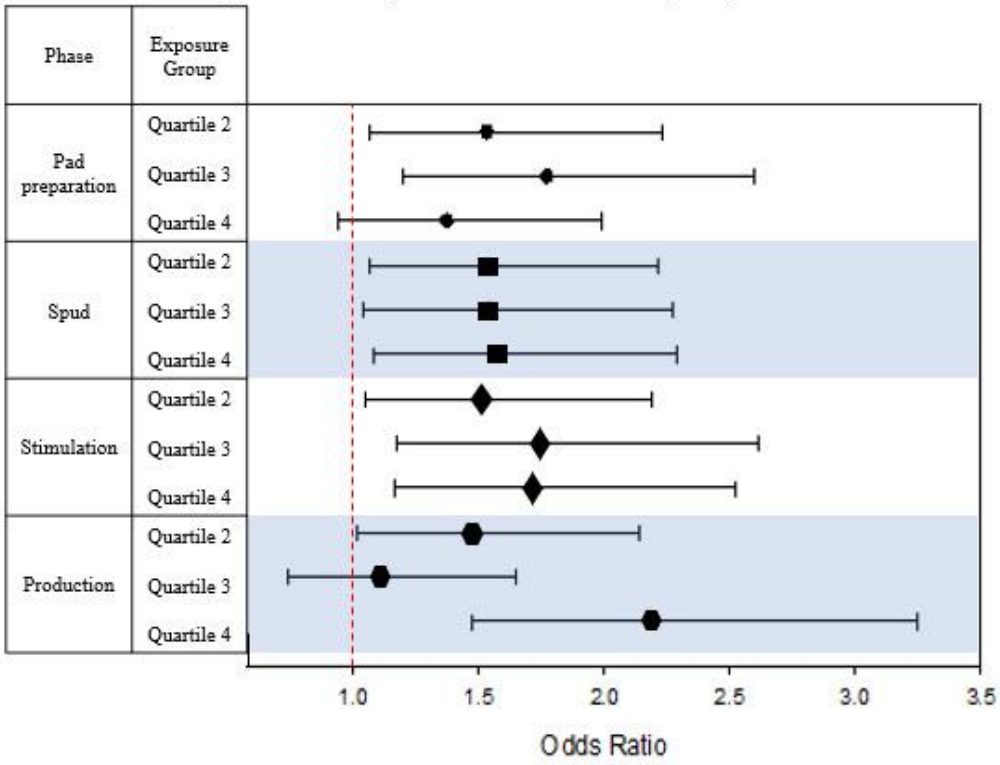
Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Oral Corticosteroid Orders



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Inverse distance weighted activity phase based on the well closest to the geocoded participant home address. (Referent: Quartile 1; Controls n = 18,693, Cases n = 13,196).

Rasmussen et al. 2016 [1,2,3]

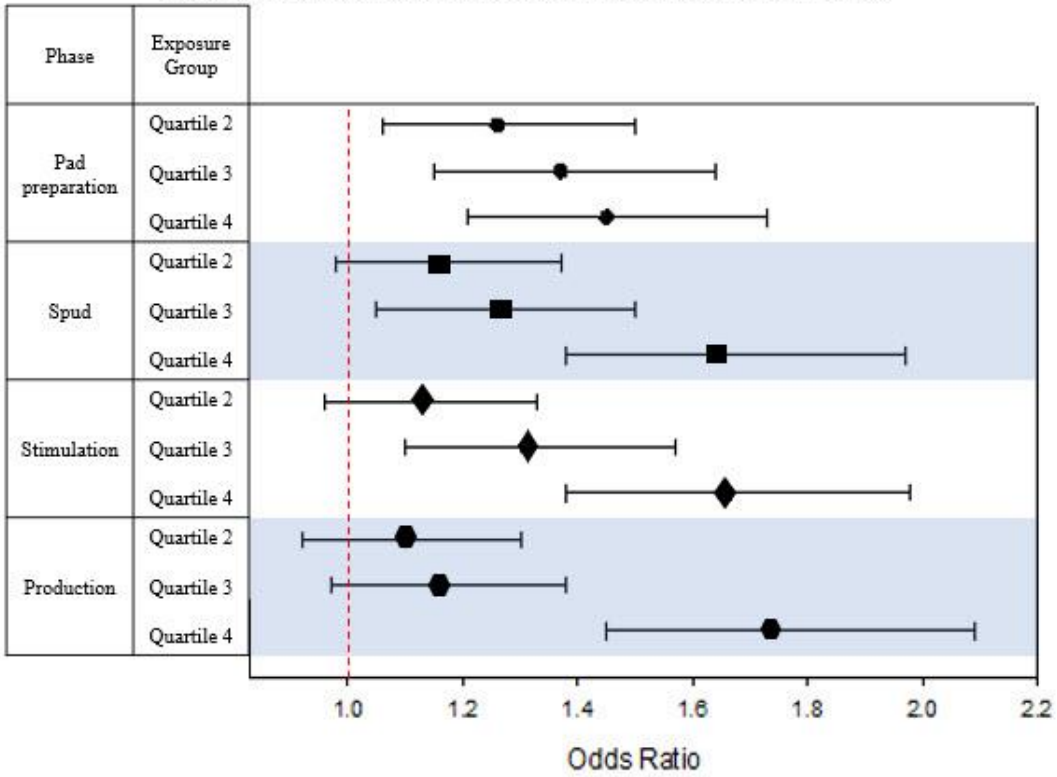
Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Asthma Emergency Room Visits



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Inverse distance weighted activity phase based on the well closest to the geocoded participant home address. (Referent: Quartile 1; Controls n = 9,350, Cases n = 1,454).

Rasmussen et al. 2016 [1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Asthma Hospitalizations



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Inverse distance weighted activity phase based on the well closest to the geocoded participant home address. (Referent: Quartile 1; Controls n = 14,104, Cases n = 3,576).

Rasmussen et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Included all asthma records from study area.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on missing data, age, residence outside of study area, and respiratory diseases other than asthma	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to case-control study design.	
	Control group appropriate to address study question	-Clearly delineated between cases (patients with asthma who had an exacerbation event) and controls (patients with asthma who did not have an exacerbation event). -Selected controls from study population.	
	Same population over study period		No formal assessment of this assumption for the 8-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: SES, lifestyle factors, and demographic factors.	-Family history of asthma more prevalent among cases than controls across all outcomes. -Percent White lower in emergency department (ED) cases than controls. -Percent living in city higher in ED cases than controls.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained using ICD-9 diagnosis coding	ICD-9 diagnosis coding may capture other acute ailments by patients with asthma
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	-Dependent on the quality of underlying data from Pennsylvania Department of Environmental Protection, Pennsylvania Department of Conservation and Natural Resources, and US Department of Agriculture. -Used crowdsourcing to identify well pad location. -Geocoded to home address.	-Assumed residential stability throughout study period. -No discussion of quality of data.

Rasmussen et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
		-Noted that 80% remained at same address (evaluated from addresses of same study sample in previous study three-years prior to present study)	
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No
	Assess exposure in a way that addresses review question.	-Incorporated information about proximity to and number of wells. -Included only UOGD wells. -Assigned at daily-resolution.	-Did not evaluate how exposure surrogate relates to overall UOGD activity. -Included period before rapid onset of UOGD.
	Study period sufficient to capture exposure variability	Not applicable to case-control study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	Tested for cut point bias by creating continuous metrics	Limited spatial overlap between study population and UOGD wells.
	Differentiates among UOGD and its various phases	Reflected pad preparation, spud, stimulation, and production phases.	-Imputed 35% of stimulation dates -Pad development and drilling dates and duration estimated.
	Differentiates between active and non-active wells	Included only active UOGD wells study area.	
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Exposure assigned to each case and control using a 1-day lag before asthma event and tested alternative lags.	
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	-Used electronic health record data to collect covariate information. -Cases matched to controls on age, sex, and year of outcome observation	-Lifestyle factors recorded on health record not confirmed.
	Controlled for baseline conditions	-Controlled for co-morbidities and family history of asthma -Controlled for community-level factors	No control of detailed lifestyle or more detailed SES factors.
	Controlled for background exposures	-Reliable method of calculating distance to major roadway. -Controlled for temperature on day before outcome observation	No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, conventional wells that vary daily.
	Assessed time trends		No.
Analytical Methods	Analytical methods appropriate for study design	-Yes. -Described model-building procedure. -Analytical methods accounted	

Rasmussen et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
		for spatial and intra-individual correlation.	
	Report measures of precision and variability	-Presented variability of characteristics by case status. -95% Confidence intervals for odds ratios.	
	Report which statistical tests were used	-Odds ratios for multilevel logistic regression. -$p<0.05$ for significance testing.	
	Perform analysis to test sensitivity of results to alternative specifications.	-Tested for cut point bias of exposure groups -Assessed potential for residual confounding using negative control -Tested impact of unbalanced numbers of cases and controls -Tested impact of exposure misclassification by including only addresses geocoded to the home	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Discussion adequately addresses study limitations	Discussed potential for selection bias, residential mobility, limitations of exposure assessment	-No discussion of potential for residual confounding. -No discussion of hydrology or meteorology on potential exposure misclassification.
	Appropriate and complete interpretation of results	-Interpreted reported effect estimates appropriately. -Discussed potential for occupational source of exposure.	-No discussion of other potential environmental sources or explanations for observed associations.

PENG ET AL. 2018

Research Objective. The objective of this study was to assess the association between UOGD and hospitalization prevalence rates for outcomes known to be associated with exposure to air pollution.

Study Period and Location. Investigators collected county-level health and well location data for the years 2001 and 2013 for the state of Pennsylvania.

Study Population. The study population included the statewide population older than four years of age.

Outcome Ascertainment. The investigators ascertained inpatient hospital admission records for the study period from the Pennsylvania Healthcare Cost Containment Council. Specific ICD-9 codes were extracted from all inpatient admission records for acute myocardial infarction, chronic obstructive pulmonary disease, asthma, pneumonia, and upper respiratory infections.

Exposure Assessment. The investigators obtained natural gas well spud date, location, operator, and configuration data from the Pennsylvania Department of Environmental Protection's Oil and Gas reports. They obtained well-specific annual gas production from the department's well production database to identify active wells. The study investigators created four exposure metrics: (1) a binary indicator of whether there was an active well in the county in the year of the hospitalization, (2) a binary indicator of whether there was an active well in the county in the year before the hospitalization, (3) natural gas output from active wells in the county in the year of the hospitalizations, and (4) first lag of the natural gas output from active wells in the county in the year of the hospitalizations.

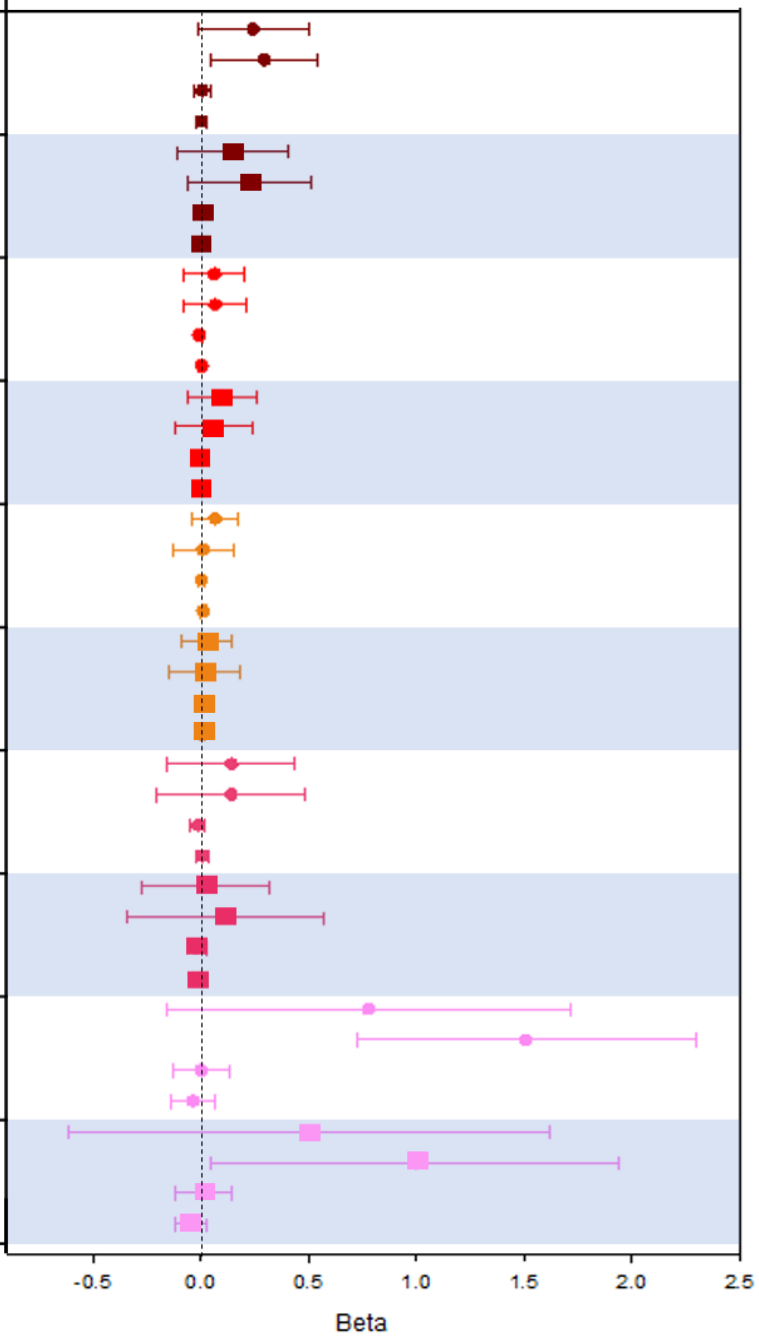
Analytical Methods and Covariate Inclusion. The investigators tested their hypothesis using a difference-in-differences model that clustered standard errors at the county level. Models were presented with and without controlling for county-specific linear trends. All models included county-level demographic variables (unemployment rate, poverty rate, median household income, population density, percentage of population in each age category, proportion female, race/ethnicity, insurance type, admission type, and the Charlson Comorbidity Index). These models also controlled for environmental covariates, including county-level annual number and total production output of new conventional wells as well as surface and underground coal production. The authors also performed a synthetic control method, assigning counties as "treated" during and after the first spud date in a given county.

Results. The following figure summarizes results as presented by the study investigators (excluding any results provided in supplementary information).

Peng et al. 2018 [1,2,3]

Betas and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Pneumonia

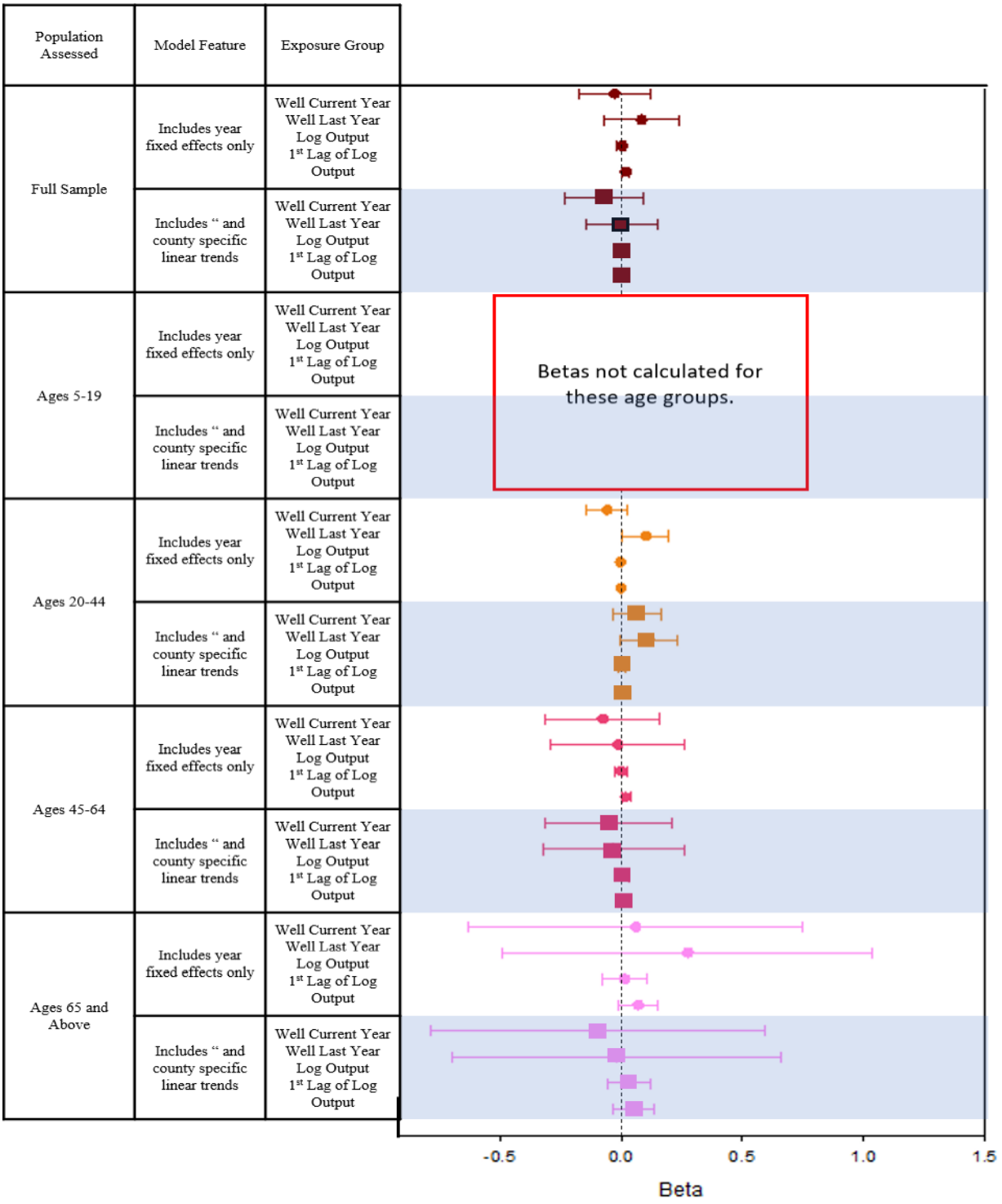
Population Assessed	Model Feature	Exposure Group
Full Sample	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 5-19	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 20-44	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 45-64	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 65 and Above	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output



1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents an outcome, and each shape represents a different model.
3. Exposure surrogate: Before or after earliest spud date and log of natural gas output from active unconventional wells within the county of residence. Total number of observations = 804.

Peng et al. 2018 [1,2,3]

Betas and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Acute Myocardial Infarction

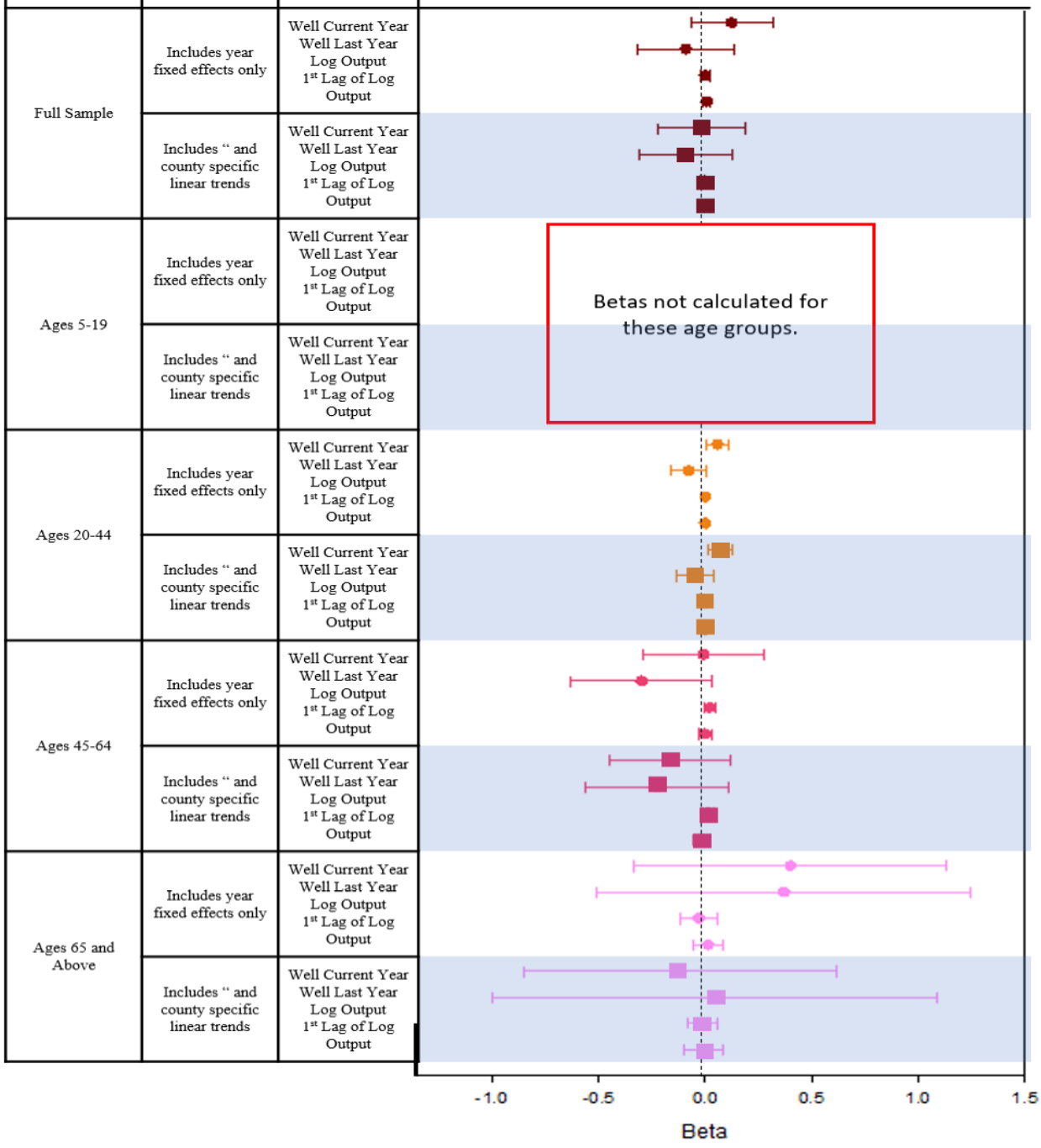


1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents an outcome, and each shape represents a different model.
3. Exposure surrogate: Before or after earliest spud date and log of natural gas output from active unconventional wells within the county of residence. Total number of observations = 804.

Peng et al. 2018 [1,2,3]

Betas and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Chronic Obstructive Pulmonary Disease

Population Assessed	Model Feature	Exposure Group
Full Sample	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 5-19	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 20-44	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 45-64	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 65 and Above	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output

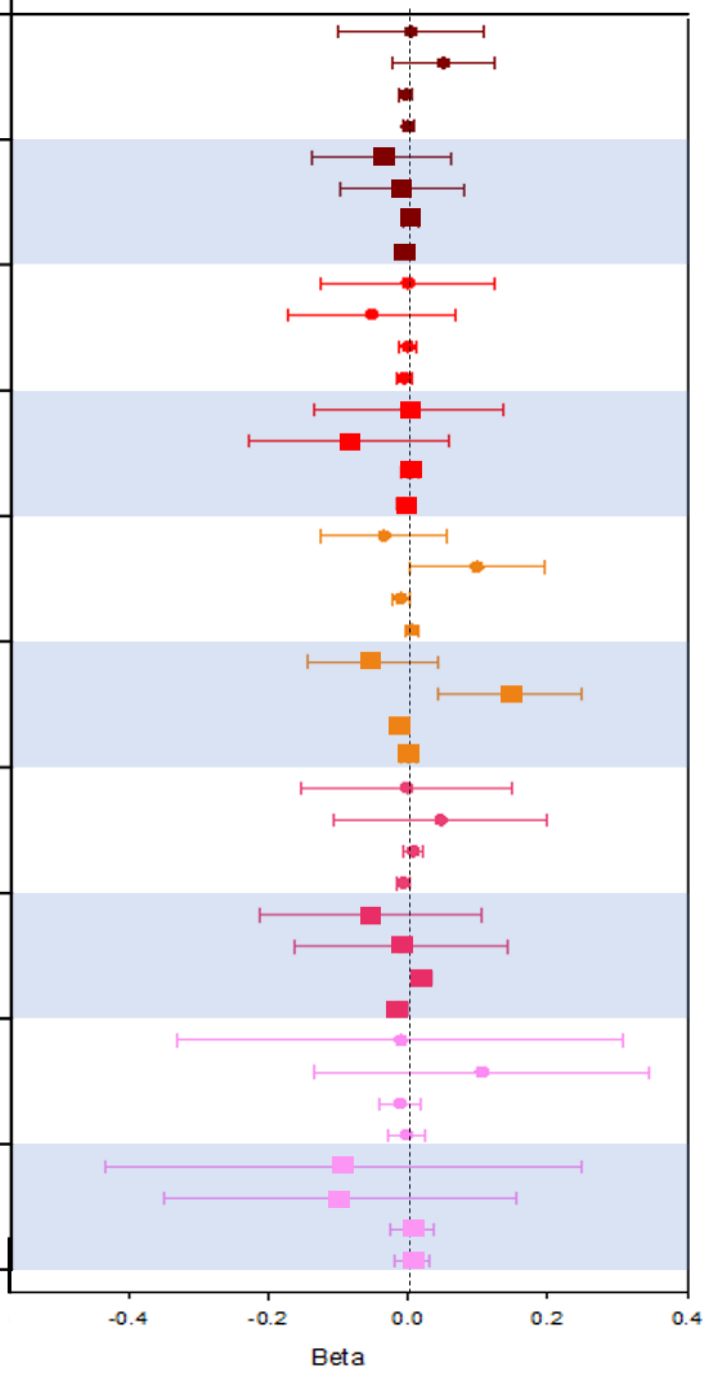


1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents an outcome, and each shape represents a different model.
3. Exposure surrogate: Before or after earliest spud date and log of natural gas output from active unconventional wells within the county of residence. Total number of observations = 804.

Peng et al. 2018 [1,2,3]

Betas and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Asthma

Population Assessed	Model Feature	Exposure Group
Full Sample	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 5-19	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 20-44	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 45-64	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 65 and Above	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output

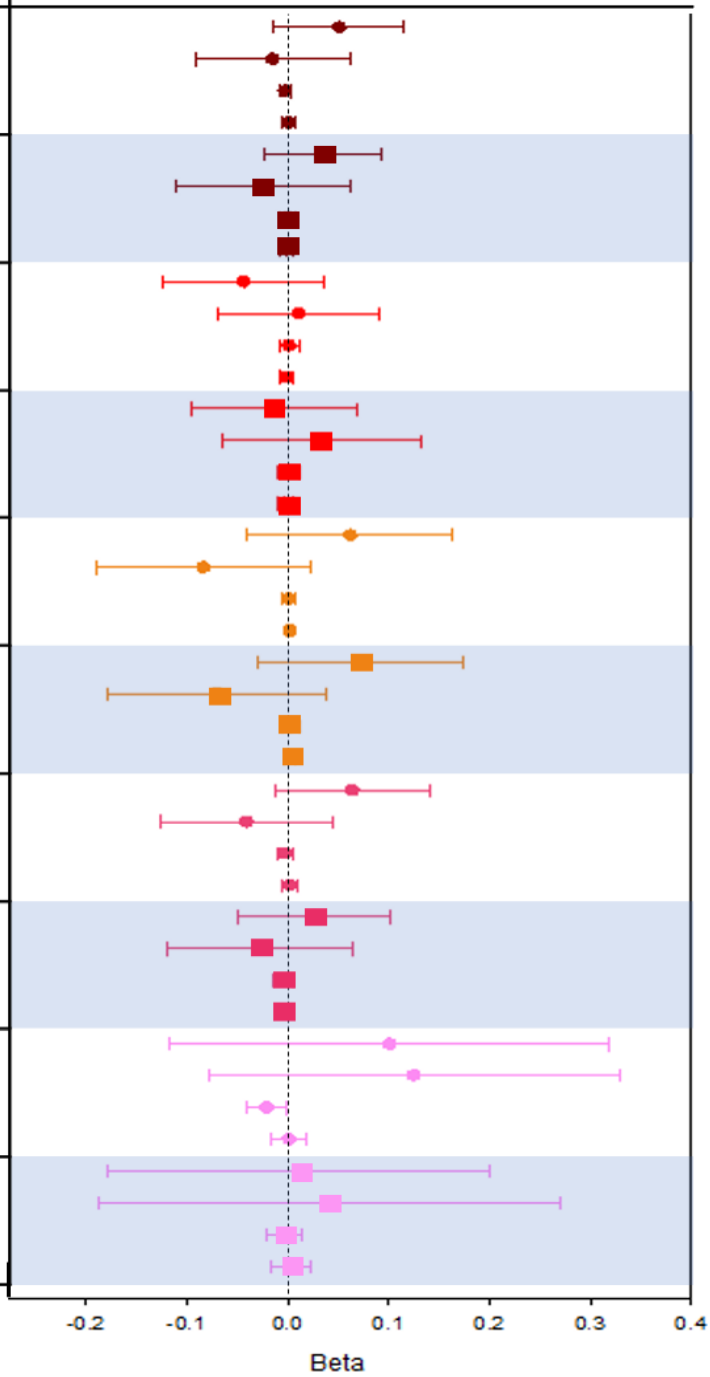


1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents an outcome, and each shape represents a different model.
3. Exposure surrogate: Before or after earliest spud date and log of natural gas output from active unconventional wells within the county of residence. Total number of observations = 804.

Peng et al. 2018 [1,2,3]

Betas and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Upper Respiratory Infections

Population Assessed	Model Feature	Exposure Group
Full Sample	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 5-19	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 20-44	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 45-64	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output
Ages 65 and Above	Includes year fixed effects only	Well Current Year Well Last Year Log Output 1st Lag of Log Output
	Includes " and county specific linear trends	Well Current Year Well Last Year Log Output 1st Lag of Log Output



1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents an outcome, and each shape represents a different model.
3. Exposure surrogate: Before or after earliest spud date and log of natural gas output from active unconventional wells within the county of residence. Total number of observations = 804.

Peng et al. 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Included all hospitalization records from study area.	
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on age	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to ecologic study design.	
	Control group appropriate to address study question	Not applicable to ecologic study design.	
	Same population over study period	Assessed this assumption using difference-in-differences analysis.	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: county-level unemployment rate, population density.	Poverty rate, coal production, number of conventional wells and conventional well output higher in counties with UOGD wells compared to counties without UOGD wells. Household median income higher in counties without UOGD wells.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained using ICD-9 diagnosis coding.	
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Dependent on the quality of underlying data from Pennsylvania Department of Environmental Protection.	-Assumed residential stability throughout study period. -No discussion of quality of data.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Used information about annual natural gas output. -Included only unconventional wells.	No consideration of potential magnitude of exposure.
	Study period sufficient to capture exposure variability	13-year study period sufficient to capture variability over time.	
	Selection of exposure groups that represent the full range of variability in UOGD.		Counties categorized as "unexposed" may include individuals that may be in

Peng et al. 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
			close proximity to UOGD in adjacent counties.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells		May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		-Exposure assigned using one-lag may not be appropriate for all outcomes assessed. -Exposure assigned during year of hospitalization (not date).
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	-Used multiple data sources to obtain county-level socioeconomic and exposure covariates, and electronic health records to obtain individual-level sex and race/ethnicity, aggregated to the county.	
	Controlled for baseline conditions	- Controlled for several county-level covariates: comorbidity index, SES factors, demographic factors, hospitalization type, and population density.	-Study type prevents control of individual-level factors. -No control of lifestyle factors.
	Controlled for background exposures	Controlled for county-level coal production, intensity and number of conventional wells.	No control of other potential environmental sources, including occupational exposures, or traffic.
	Assessed time trends	-Used difference-in-differences method. -Controlled for county-specific linear trends. -Visually assessed trends using synthetic control method.	
Analytical Methods	Analytical methods appropriate for study design	-Yes. -Analytical methods accounted for spatial and intra-county correlation.	Did not describe covariate selection into model.
	Report measures of precision and variability	-Presented variability of characteristics by county-level exposure groups. -Standard errors for coefficients.	
	Report which statistical tests were used	-Beta coefficients for difference-in-differences. -$p<0.1$, 0.05, and 0.01 for significance testing. -Calculated the family-wise error rate adjusted p values.	

Peng et al. 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Perform analysis to test sensitivity of results to alternative specifications.	-Assessed confounding by excluding urban counties -Tested alternative exposure specifications -Tested for spurious findings.	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Discussion adequately addresses study limitations	Discussed residential mobility, temporal misalignment of exposure and outcome, biological relevance.	No discussion of limitations of ecologic study design.
	Appropriate and complete interpretation of results		No discussion of other potential explanations for observed associations.

Research Objective. The objective of this study was to quantify the association between unconventional natural gas development and pediatric asthma hospitalizations.

Study Period and Location. This study was conducted in ZIP codes within counties that overlap the Marcellus Shale areas in Pennsylvania. Data were collected for the years 2003 to 2014.

Study Population. The study population included patients between 2 and 18 years of age living in rural Pennsylvania counties on the Marcellus Shale who were hospitalized because of asthma (n = 15,837). Residents of non-rural counties were excluded.

Outcome Ascertainment. The investigators obtained data on inpatient hospitalization from the Pennsylvania Healthcare Cost Containment Council. The analysis included patients with a 493.X ICD-9 code, acute asthma exacerbation. Investigators categorized the outcome as binary for the ZIP code, quarter, and year of hospitalization.

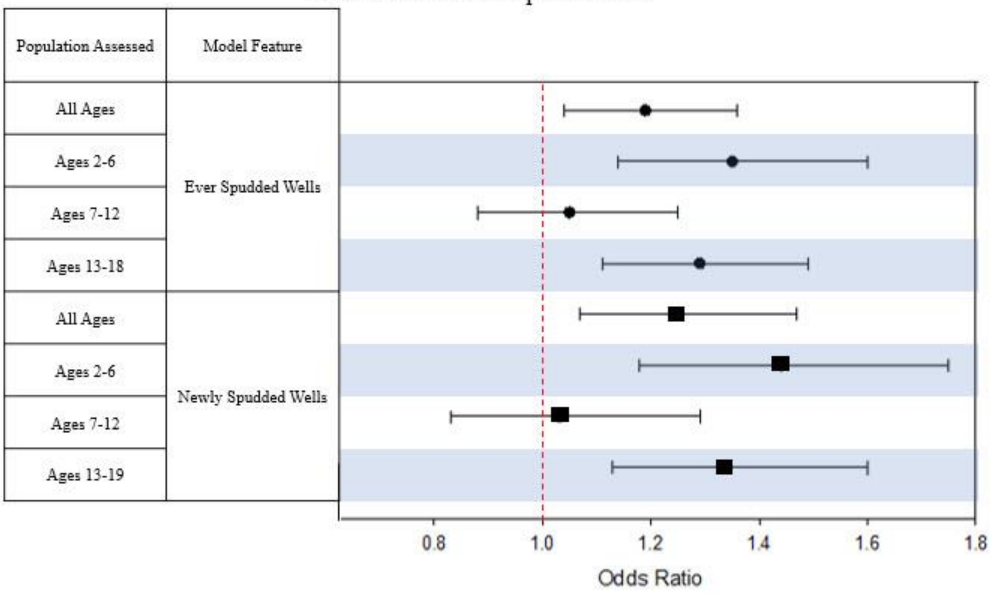
Exposure Assessment. The investigators obtained well data from the Pennsylvania Department of Environmental Protection. The investigators created three ZIP-code level exposure surrogates for each quarter and year of the study period: a binary contemporaneous variable for newly spudded wells, a binary cumulative variable for ever-spudded wells, and tertiles of cumulative count of ever-drilled wells.

Analytical Methods and Covariate Inclusion. Investigators fit mixed-effects logistic regression models with a random intercept for ZIP code and fixed effects for year and quarter. Investigators included the following covariates in analytical models: ZIP-code level proportion of hospitalizations for sex, race, ethnicity, and insurance type; ZIP code-level population density and background respiratory hazard index from the National Air Toxics Assessment; and county-level unemployment, childhood poverty, and median household income. Investigators also stratified analyses by age group. In sensitivity analyses, the investigators included conventional oil and gas development in models, ran conditional likelihood logistic regression models, included a continuous variable for UNGD well count, and restricted analysis to ZIP codes in the highest and lowest quintiles of emissions. Investigators also tested whether the UNGD surrogate represented air emissions by regressing sum of annual ZIP-code level emissions on the UNGD surrogate.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

Willis et al. 2018 [1,2,3]

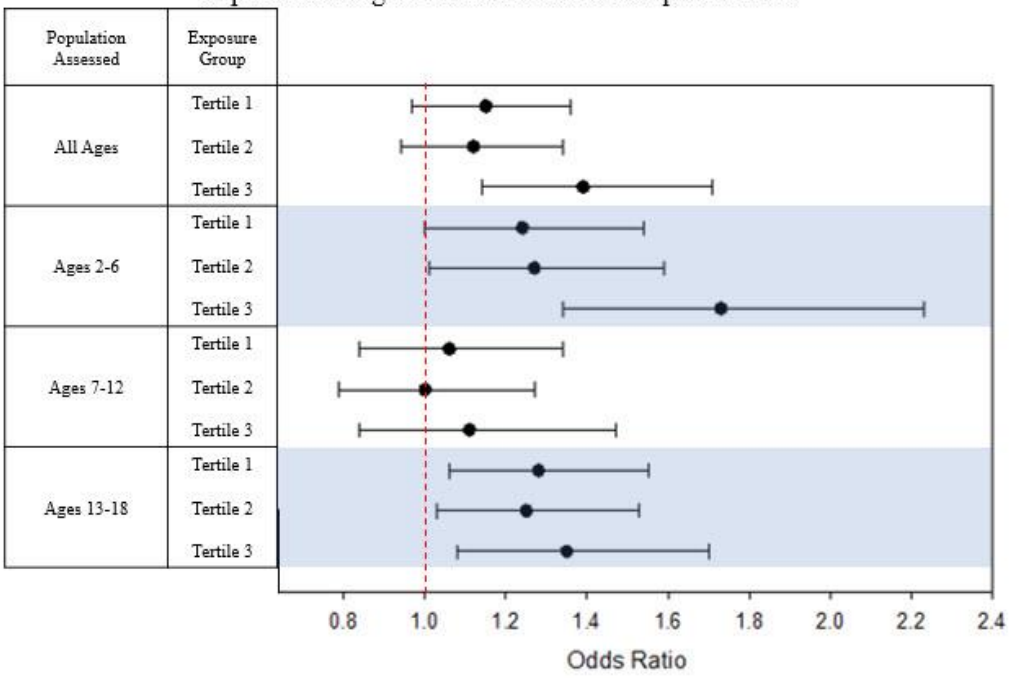
Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Pediatric Asthma Hospitalizations



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Well count within participant ZIP code of ever or newly spudded wells in quarter and year of hospitalization (binary variable). Unexposed n = 14,767, Exposed n = 1,070. Referent: No ever or newly spudded wells in ZIP code.

Willis et al. 2018 [1,2]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate Pediatric Asthma Hospitalizations



1. Plotted from results as presented in the study (excluding supplemental information).
2. Exposure Surrogate: Count of ever spudded wells in quarter and year of hospitalization (tertiles). Unexposed n = 14,767, exposed n = 1,070. Referent: No ever-spudded wells in ZIP code.

Willis et al., 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Study population representative of rural PA regions.	
	Inclusion/Exclusion criteria specified	-Detailed discussions of the criteria used to collect study sample. -Exclusions based on residence in an urban county.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to retrospective study design.	
	Control group appropriate to address study question	Not applicable to retrospective cohort study design.	
	Same population over study period		Did not assess this assumption for 12-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: age, proportion female, race/ethnicity, income, insurance type, time of year of hospitalization, respiratory hazard index, poverty and unemployment rates	Population density and emergency hospitalizations higher in ZIP-codes with wells drilled contemporaneously with asthma hospitalizations compared to ZIP-codes without.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Ascertained using ICD-9 diagnosis coding.	ICD-9 coding may not capture all asthma exacerbations.
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Reliability depends on the quality of data available at Pennsylvania Department of Environmental Quality.	-Assumed residential stability throughout study period. -Pennsylvania Unconventional Natural Gas Emission Inventory data collected via self-report.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Used tons of yearly emissions for each UNGD site -Regressed exposure surrogate on annual emissions of multiple different chemicals to	-Included period before rapid onset of UNGD. -Exposure assigned based on year of hospitalization. -ZIP-code level analysis.

Willis et al., 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
		test what exposure surrogates represent.	-Emissions data available for four years of the 11-year study period.
	Study period sufficient to capture exposure variability	Not applicable to retrospective study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	Tested for cut point bias using both continuous and categorical exposure surrogates.	
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included UNGD wells post-spud date.	May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		-Exposure is assessed at the quarter or annual level, which is not appropriate for the outcomes assessed. -Exposure may not overlap with outcome time-at-risk.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data to collect covariate information on the individual and ZIP code or county level.	
	Controlled for baseline conditions	-Controlled for basic sociodemographic characteristics at the community level: annual proportions of hospitalizations by sex, race, ethnicity, and insurance, and SES measures.	No control of individual-level factors, including lifestyle, co-morbidities, or detailed SES factors.
	Controlled for background exposures	-Controlled for an index quantifying non-UNGD respiratory hazards. -Controlled for non-UNGD wells in sensitivity analysis. -Included population residing only in rural ZIP codes.	No control of other potential environmental sources, including industrial sources, or individual co-exposures that vary daily.
	Assessed time trends	Controlled for hospitalization proportions over time, year, and quarter of outcome observation.	No formal assessment of impact of time trend over 11-year study period.

Willis et al., 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Analytical Methods	Analytical methods appropriate for study design	Accounted for spatial and temporal correlation in model	-Investigators did not use Poisson models, which is more commonly used for count data. -Unclear how analytical models handles ZIP-codes with small numbers of hospitalizations -Did not describe model-building procedure.
	Report measures of precision and variability	95% Confidence intervals for odds ratios.	
	Report which statistical tests were used	Odds ratio for logistic regression.	
	Perform analysis to test sensitivity of results to alternative specifications.	-Assessed effect modification by age subgroups. -Tested impact of random intercepts. -Tested cumulative well count. -Restricted analysis to ZIP codes with the lowest and highest annual emissions.	
Results and Discussion	All findings reported for analyses described in paper		-Tested associations for 180 pollutants but presented results for only 16 -Difference-in-differences results not presented.
	Discussion adequately addresses study limitations	Discuss potential for exposure misclassification and lack of individual-level data	-No discussion of within-ZIP-code correlations and the non-independence of the individual-pollutant analyses. -No discussion of potential exposure misclassification -No discussion of low temporal resolution of exposure surrogate. -No discussion of meteorology or hydrology.
	Appropriate and complete interpretation of results		-Unclear description of methods and interpretation of results. -Do not discuss disparate findings between age groups or lack of dose-response. -Interpretation does not consider pre-drill temporal trends. -Interpretation does not consider multiple hypothesis testing. -Did not fully describe their difference-in-differences analysis.

McKenzie et al. 2019

Research Objective. The objective of this study was to evaluate the association between indicators of cardiovascular disease and oil and gas development activity in Northeastern Colorado.

Study Period and Location. This study was conducted in the municipalities of Fort Collins, Windsor, and Greeley in Colorado. Data were collected between October 2015 and May 2016.

Study Population. The study population included 97 men (n = 28) and women (n = 69), who were ≥ 18 years of age. Investigators excluded those who were pregnant; tobacco or marijuana smokers; used anti-inflammatory medication, were occupationally exposed to dust, fumes, solvents, or oil and gas activities; or had a history of chronic inflammatory diseases. All participants resided full-time in Fort Collins, Windsor, or Greeley, CO.

Outcome Ascertainment. The investigators evaluated measures of cardiovascular health and systemic inflammation. Measures of cardiovascular health included augmentation index and systolic and diastolic blood pressure. Measures of systemic inflammation included interleukin-1β (IL-1β), interleukin-6 (IL-6), interleukin-8 (IL-8), and tumor necrosis factor–alpha (TNF-α). Measurements were made over the course of three study visits. Eighty percent of study participants completed all three study visits.

Exposure Assessment. The investigators geocoded each participant's residential address and obtained the latitude and longitude coordinates for oil and gas wells from the Colorado Oil and Gas Information System. Investigators calculated the distance between residential address and wells within a 16-km radius of each home. The investigators then applied an intensity adjusted–inverse distance weighted (IA-IDW) model, calculated over the 9-month study period. The surrogate incorporated well-specific information about location, number of wells, activity phase, use of green completion, production volume, number of tanks on well pad, and an intensity factor that represented estimated emission rates of select VOCs by phase. The final exposure surrogate was divided into tertiles, with the first tertile designated as the referent group.

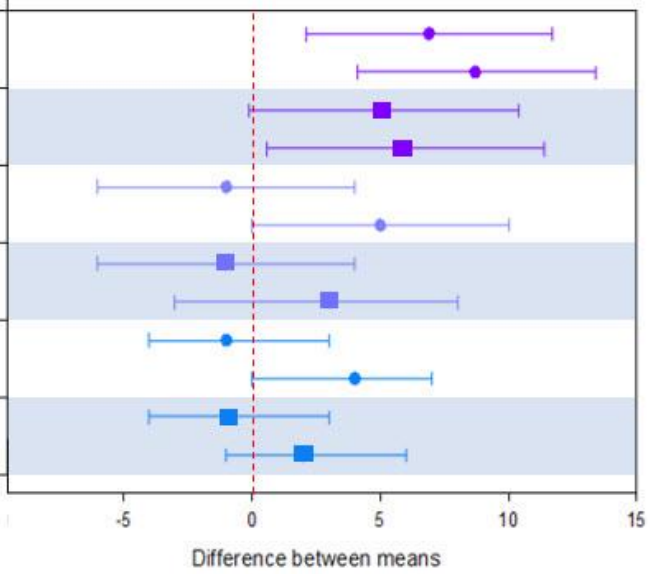
Analytical Methods and Covariate Inclusion. The investigators fit linear mixed models with random intercepts for each participant. All models were adjusted for age, sex, race/ethnicity, body mass index, educational attainment, income, and employment status. The investigators evaluated the final models for residual spatial autocorrelation using semivariograms.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

McKenzie et al. 2019 [1,2,3]

Difference Between Means and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Cardiovascular Health Endpoints

End Point	Model Feature	Exposure Group
Augmentation Index (% at HR 75)	Adjusted for age and sex	Tertile 2
		Tertile 3
	Adjusted for age, sex, race/ethnicity, BMI, education, income, and employment	Tertile 2
		Tertile 3
Systolic Blood Pressure (mm Hg)	Adjusted for age and sex	Tertile 2
		Tertile 3
	Adjusted for age, sex, race/ethnicity, BMI, education, income, and employment	Tertile 2
		Tertile 3
Diastolic Blood Pressure (mm Hg)	Adjusted for age and sex	Tertile 2
		Tertile 3
	Adjusted for age, sex, race/ethnicity, BMI, education, income, and employment	Tertile 2
		Tertile 3

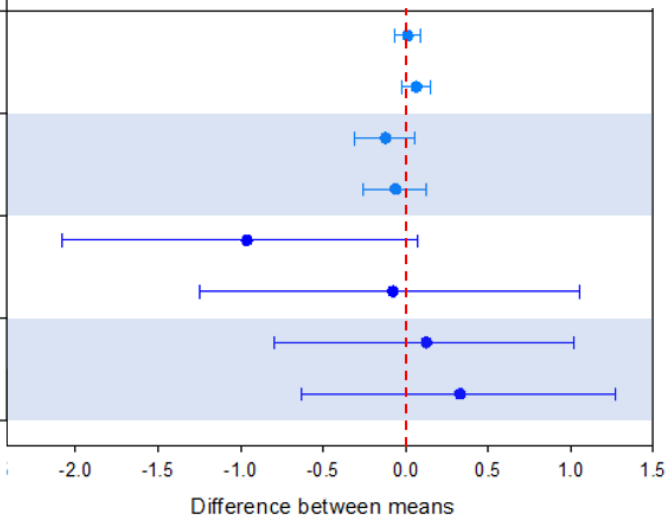


1. Plotted from results as presented in the study (excluding supplemental information).
2. Each color represents a different outcome, and each shape represents a different model.
3. Exposure Surrogate: Spatiotemporal activity model within 16 kilometers of residence. (Referent: Tertile 1; n = 97).

McKenzie et al. 2019 [1,2,3,4]

Difference Between Means and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Inflammation Markers

End Point	Exposure Group
Interleukin-1β	Tertile 2
	Tertile 3
Interleukin-6	Tertile 2
	Tertile 3
Interleukin-8	Tertile 2
	Tertile 3
Tumor Necrosis Factor	Tertile 2
	Tertile 3



1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure Surrogate: Spatiotemporal activity model within 16 kilometers of residence. (Referent: Tertile 1; n = 97).
4. Results adjusted for age, sex, race/ethnicity, BMI, education, income, and employment status

McKenzie et al. 2019: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population		Did not describe the intended study population.
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on lifestyle factors, medication use, co-morbidities, occupational exposure to chemicals, and residence.	No discussion of study participant recruitment or selection methods.
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Participation across three visits similar among exposure groups.	
	Control group appropriate to address study question	Not applicable to cross-sectional study design.	
	Same population over study period	Not applicable to cross-sectional study design.	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: visit participation, body mass index, race/ethnicity, and sex.	Third tertiles includes higher percent of: Greeley and Windsor residents, unemployment, older age, and higher income participants compared to referent.
Outcome Assessment	Outcome ascertained using valid and reliable measures	-Valid and reliable measures of systolic and diastolic blood pressure (Beevers et al., 2001) -Valid and reliable collection and analysis of inflammatory markers in blood (Vasunilashorn et al., 2015)	Did not report whether augmentation index was collected by trained professionals using valid methods.
	Outcome assessors blinded to exposure status	Inflammatory markers analyst blind to exposure status.	Did not report whether analysts measuring blood pressure and augmentation index were blind to exposure status.
	No systematic differences in outcome ascertainment or reporting between exposure groups	Ascertained identically in all exposure groups.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	-Dependent on quality of underlying data (COGIS). -Rooftop geocoding.	-Assumed residential stability throughout study period. -No discussion of quality of COGIS.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No

McKenzie et al. 2019: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Assess exposure in a way that addresses review question.	-Incorporated information about proximity, number of wells, activity phase, production volume, whether green completion was used, and the number of tanks on a well pad, and an intensity factor that represents estimated emission rates of select VOCs. -In a separate study, exposure surrogate was evaluated against emissions data collected from one of the two towns included in this study.	Did not differentiate between unconventional and conventional wells
	Study period sufficient to capture exposure variability	Not applicable to cross-sectional study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.		Did not test for cut point bias.
	Differentiates among UOGD and its various phases	Incorporated temporal component to estimate exposure during drilling, hydraulic fracturing and flowback, and production phases.	Phase duration estimated.
	Differentiates between active and non-active wells	Included all active oil and gas wells reported in Colorado Oil and Gas Information System (COGIS) from August 2015 to April 2016.	
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		Did not specify whether 2 month averaging period before sample is sufficient for outcomes studied.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used surveys to collect lifestyle factors, medical histories, and demographic information for all exposure groups.	
	Controlled for baseline conditions	Controlled for age, sex, race/ethnicity, BMI, educational attainment, income, and employment status.	No assessment of potential for confounding by community-level factors.
	Controlled for background exposures	Collected information on occupational exposure to dust, fumes, solvents and oil and gas activity.	No control of other potential environmental sources, including industrial sources, traffic, or conventional wells.
	Assessed time trends	Not applicable to cross-sectional study	
Analytical Methods	Analytical methods appropriate for study design	Analytical method accounted for intra-individual correlation of outcome measures.	Did not describe model-building procedure.

McKenzie et al. 2019: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Report measures of precision and variability	-Presented variability of characteristics by exposure group. -95% Confidence intervals for odds ratios.	
	Report which statistical tests were used	Yes.	
	Perform analysis to test sensitivity of results to alternative specifications.	-Evaluated interactions of exposure and participant sex, age, self-report stress, physical activity, exposure to VOCs, and food or drink consumption prior to outcome measure. -Performed sensitivity analysis on subsets of participants. -Assessed the potential for spatial autocorrelation.	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Discussion adequately addresses study limitations	Mentioned small sample size, potential for residual confounding, cross-sectional study design, and lack of direct noise and air pollution measures.	Did not discuss zero participants in referent group from Greeley or in the high group from Fort Collins.
	Appropriate and complete interpretation of results	-Interpreted reported effect estimates appropriately. -Discussed biological plausibility and clinical significance.	No discussion of other potential environmental sources or explanations for observed associations.
Beevers et al. 2001 *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1120141/* Vasunilashorn et al. 2015 *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4817082/*			

RABINOWITZ ET AL. 2016

Research Objective. The objective of this study was to determine whether residential proximity to UOGD was associated with a variety of self-reported health symptoms.

Study Period and Location. This cross-sectional study took place in Washington County in Southwestern Pennsylvania. The survey was administered in summer 2012.

Study Population. The study population consisted of 180 households (492 individuals) located in Washington County (of 760 initially selected) with ground-fed water wells.

Outcome Ascertainment. Study personnel administered an in-person survey to collect self-reported dermal, upper and lower respiratory, cardiac, gastrointestinal, and neurologic symptoms. Study participants were asked whether they or other household members had experienced any of the listed symptoms in the past year.

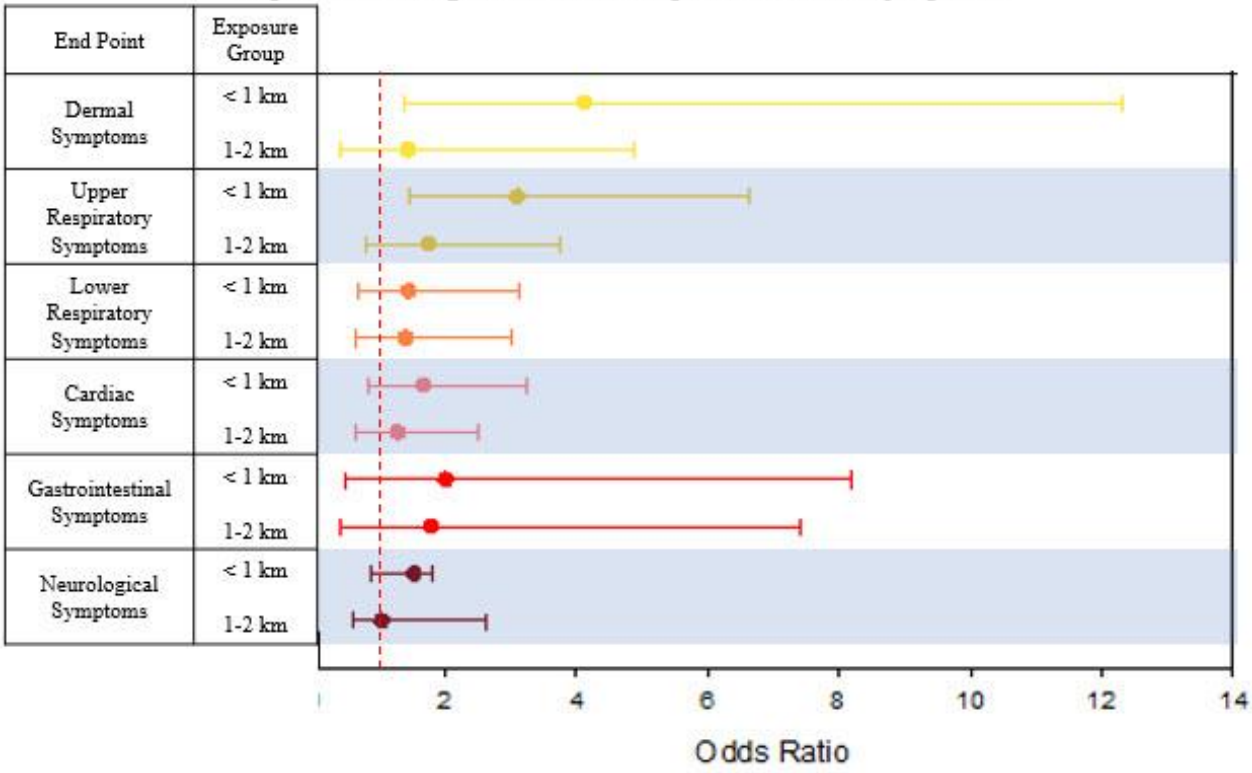
Exposure Assessment. At the time of survey, study personnel recorded global positioning system coordinates of each participating home. Gas well permit data were obtained from the Pennsylvania Spatial Data Access program (Penn State University), which provided data on well spud date, coordinates, age and type. Distance between survey household coordinates and distance to the nearest "active natural gas" well was calculated using ArcGIS and categorized into three groups: <1 km, 1–2 km, and >2 km. Of 624 active natural gas wells, 95% were horizontally drilled and 95% had spud dates from 2008–2012.

Analytical Methods and Covariate Inclusion. Descriptive statistics were performed using chi-square and ANOVA. Generalized linear mixed models, with a random effect for household, were used to assess the association between the exposure surrogate and reported symptoms. The model included the following covariates: age of respondent, sex, average adult household education, smoker in household, awareness of environmental hazard nearby, employment type (blue collar, office services, management/professional, unemployed), and presence of animals in the home.

Results. The following figure summarizes results as presented by the study investigators (excluding any results provided in supplementary information).

Rabinowitz et al. 2015 [1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Self-Reported Health Symptoms



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure Surrogate: Whether nearest well to household is <1 kilometer, 1-2 kilometers, or >2 kilometers. (Referent: >2 kilometers from closest well; n = 492).

Rabinowitz et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Random sampling and household selection approach suggests representative population	
	Inclusion/Exclusion criteria specified	-Population selection process described in detail. -Exclusions based on accessibility of home to road, non-ground-fed water supply, whether house was occupied during study period, residence in study area.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to cross-sectional study design.	
	Control group appropriate to address study question	Not applicable to cross-sectional study design.	
	Same population over study period	Data collected over one year.	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: sex, education, years in household, BMI, water source.	-Proportion of children in household and household smoking lower in highest exposure group. -Age of respondent, years since closest wells spud date, environmental risk awareness, dissatisfied with odor in environment, water with unnatural appearance, and proportion of blue-collar workers higher in highest exposure group.
Outcome Assessment	Outcome ascertained using valid and reliable measures	May capture health status of population not otherwise captured in large, administrative datasets.	Not medically confirmed and possibly subject to information bias.
	Outcome assessors blinded to exposure status	Survey team was unaware of proximity analysis.	Unclear whether there were visual clues of UOGD when traveling to households to administer survey.
	No systematic differences in outcome ascertainment or reporting between exposure groups	Controlled for awareness of environmental risk.	Some systematic differences found [e.g., refusal to participate rate higher among those living further from wells, environmental awareness associated with symptoms prevalence (though controlled for analytically)]
Exposure Assessment	Performed using valid, reliable and sensitive methods	-Dependent on the quality of underlying data from Pennsylvania Spatial Data Access. -Geocoded to home address.	-Unable to account for residential mobility throughout potentially exposed period. -No discussion of quality of data.

Rabinowitz et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.		-Did not evaluate how exposure surrogate relates to UOGD activity. -Unclear study objective: incorporated information about proximity to all natural gas wells, though background and abstract discuss potential exposures from UOGD processes.
	Study period sufficient to capture exposure variability	Not applicable to cross-sectional study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	Choice of distance <1 km supported with previous literature, assuming exposure via the water pathway.	-No justification for choice of radius within which to calculate exposure surrogate >1 km. -Did not test for cut-point bias.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included only active wells study area.	
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Investigators did not consider outcomes with long latency periods.	Unclear if one year is sufficient to develop all outcomes studied.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Collected a basic current demographic, SES, and lifestyle factors.	Unclear if differential reporting of covariates related to the exposure or outcome occurred (e.g., smoking, water appearance, water taste/odor, environmental risk awareness).
	Controlled for baseline conditions	-Controlled for individual-level occupation, basic demographic and SES characteristics. -Collected household-level smoking, years occupying household, and water appearance and taste, environmental odor.	-Lacking detailed information on co-morbidities, SES, family medical history, and lifestyle factors (except smoking). -No control of community-level factors. -No control for years in household or water-related variables
	Controlled for background exposures	Controlled for animal in household.	No control of other potential environmental sources, including occupational exposures, industrial

Rabinowitz et al. 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
			sources, traffic, or conventional wells.
	Assessed time trends	Restricted analysis to one year.	
Analytical Methods	Analytical methods appropriate for study design	-Yes. -Analytical methods accounted for intra-household correlation.	Selection of covariates for model inclusion not described.
	Report measures of precision and variability	-Presented variability of characteristics by exposure status. -95% Confidence intervals for odds ratios.	Temporal variability of exposure not presented.
	Report which statistical tests were used	-Odds ratios for multilevel linear regression. -$p<0.05$ for significance testing.	
	Perform analysis to test sensitivity of results to alternative specifications.		No.
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Discussion adequately addresses study limitations	Detailed discussion of potential for selection bias, potential recall bias, potential for residual confounding, lack of medical confirmation of outcomes, multiple comparisons, and limitations of exposure assessment	-No discussion of hydrology, meteorology, or medical outcome groupings. -Exposure pathways presented in discussion not discernable with the resolution of the data.
	Appropriate and complete interpretation of results	-Interpreted reported effect estimates appropriately. -Discussed possible alternative explanations for observed associations.	Interpretation does not consider multiple hypothesis testing.

Research Objective. The objective of this study was to examine the association between UOGD and nasal symptoms, migraine, and fatigue.

Study Period and Location. The investigators collected cross-sectional survey data in 2014 from participants living in central and northeastern Pennsylvania.

Study Population. The study population consisted of randomly selected Geisinger Health System primary care patients over 18 years of age. Of the 23,700 selected survey recipients, 7,785 Pennsylvania residents responded. Racial/ethnic minorities and individuals with higher risk of chronic rhinosinusitis were oversampled.

Outcome Ascertainment. The investigators defined chronic rhinosinusitis cases as those who reported having two or more symptoms of nasal congestion/obstruction, nasal discharge, smell loss, and facial pain or pressure at least "most of the time" in the past three months, using previously established diagnostic criteria. Migraine cases were identified using a validated migraine questionnaire, which asked about frequency and duration of headache- and migraine-associated symptoms in the past 12 months. A validated questionnaire was used that assessed the frequency of fatigue and fatigue-related disability within the past week.

Exposure Assessment. The investigators obtained well location, spud, production and stimulation dates, drilling depth, and volume of natural gas production between 2005 and 2014 from the Pennsylvania Department of Environmental Protection and the Pennsylvania Department of Conservation and Natural Resources. Crowd-sourced photos from SkyTruth were used to determine well location. The investigators developed an IDW-squared exposure surrogate for each of four UOGD phases. Investigators averaged the surrogate over the 90 days before survey administration. The four-phase metrics were *z*-transformed and summed for an overall *z*-score UOGD exposure metric. This variable was then categorized into quartiles.

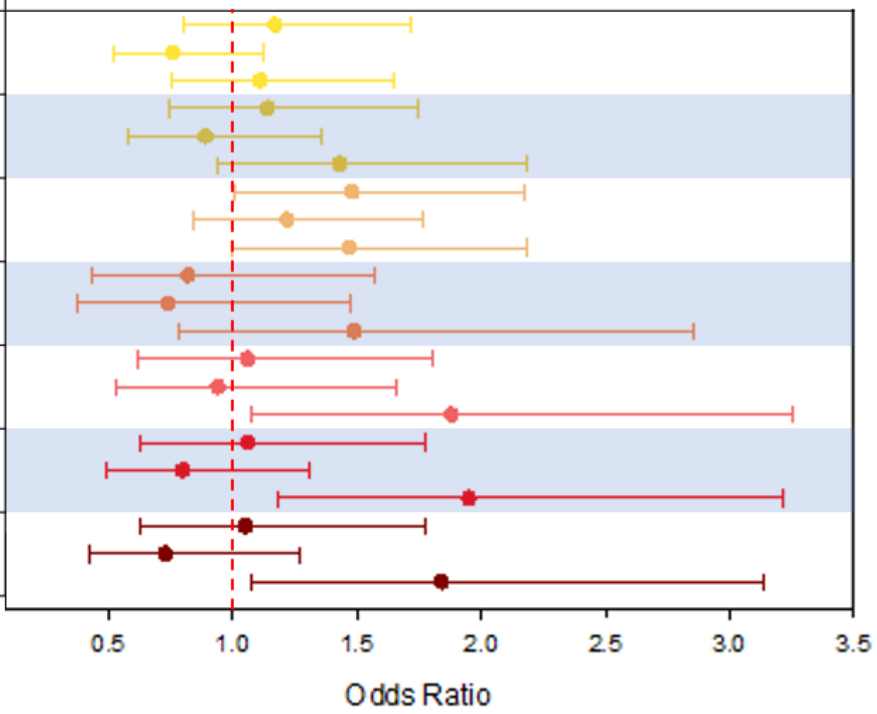
Analytical Methods and Covariate Inclusion. The investigators compared categorical variables using chi-square tests and continuous variables using *t*-tests. They used weighted logistic regression to examine the association between the UOGD surrogate and reported symptoms. The following covariates were assessed for model inclusion a priori: sex, race/ethnicity, age, body mass index, community socioeconomic deprivation, receipt of Medical Assistance (i.e., Medicaid), and smoking status. Additionally, the investigators stratified by date of symptom onset (before and after 2006) to assess symptom prevalence during periods before and after UOGD commenced in Pennsylvania.

Results. The following figure summarizes results as presented by the study investigators (excluding any results provided in supplementary information).

Tustin et al. 2016[1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Reported Chronic Rhinosinusitis, Migraine, Higher Levels of Fatigue

Outcome	Cases	Exposure Group
Current Chronic Rhinosinusitis Only	736	Quartile 2 Quartile 3 Quartile 4
Migraine Headache Only	580	Quartile 2 Quartile 3 Quartile 4
Higher Fatigue Only	666	Quartile 2 Quartile 3 Quartile 4
Current Chronic Rhinosinusitis and Migraine	266	Quartile 2 Quartile 3 Quartile 4
Current Chronic Rhinosinusitis and Higher Fatigue	347	Quartile 2 Quartile 3 Quartile 4
Migraine Headache and Higher Fatigue	420	Quartile 2 Quartile 3 Quartile 4
All Three Outcomes	496	Quartile 2 Quartile 3 Quartile 4



1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure Surrogate: Inverse distance weighted activity phase Z-square transformed in Pennsylvania. (Referent: Quartile 1; Reference group $n = 1{,}380$).

Tustin 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	-Random sampling approach for study sample selection. -Study sample representative of medical system catchment area	-Survey respondents more likely to have poorer health than non-respondents -Oversampled people with nasal and sinus symptoms and racial/ethnic minorities.
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on age, history of outcome assessed, co-morbidities, residence outside of study area, and failure to complete questionnaire.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to cross-sectional study design.	
	Control group appropriate to address study question	-Clearly delineated between cases and controls. -Selected controls from study population.	
	Same population over study period	Data collected over one year.	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: never and former smoker, BMI, residence in a borough or township.	-All outcomes, combined: Cases younger, more likely to receive Medical Assistance, have higher socioeconomic deprivation, and be a current smoking than controls. -All migraine outcomes: Higher proportions of female cases than controls; cases younger -Fatigue and CRS, Current CRS and migraine and current CRS only: Higher proportion of white non-Hispanic cases than controls.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Used validated questionnaire for migraine and fatigue.	Outcomes not medically confirmed and possibly subject to information bias.
	Outcome assessors blinded to exposure status	Study questionnaire did not mention UOGD in explanation of study objectives.	
	No systematic differences in outcome ascertainment or reporting between exposure groups		Not specified.
Exposure Assessment	Performed using valid, reliable and sensitive methods	-Dependent on the quality of underlying data from Pennsylvania Department of Environmental Protection and Pennsylvania	-Assumed residential stability throughout potentially exposed period. -No discussion of quality of data.

Tustin 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
		Department of Conservation and Natural Resources. -Used crowdsourcing to identify well pad location.	-Geocoding procedure not described.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Incorporated information about proximity to and number of wells. -Included only UOGD wells. -Assigned at daily-resolution.	No evaluation of whether exposure surrogate represents UOGD activity.
	Study period sufficient to capture exposure variability	Not applicable to case-control study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.		-Limited spatial overlap between study population and UOGD wells. -Did not test for cut-point bias.
	Differentiates among UOGD and its various phases	Reflected pad preparation, spud, stimulation, and production phases.	-Pad development and drilling dates and duration estimated. -Unable to distinguish between phases because of collinearity (aggregated phases to z-score).
	Differentiates between active and non-active wells	Included only active UOGD wells study area.	May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Exposure assigned to each case and control using a 90-day averaging period and tested 7- and 365-day averaging periods.	Unclear if averaging periods are sufficient to develop all outcomes studied.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data and administered questionnaires to collect individual-level covariate information, and Census to collect community-level data.	Lifestyle factors recorded on health record not confirmed using questionnaire.
	Controlled for baseline conditions	-Collected several co-morbidity and medication use variables, and basic individual- and community-level SES and demographic characteristics -Controlled for sex, race/ethnicity, age, and receipt of Medical Assistance.	Lacking detailed information on SES, family medical history, and lifestyle factors (except smoking).

Tustin 2016: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Controlled for background exposures	Controlled for smoking status.	No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells.
	Assessed time trends		No formal assessment of this assumption for the 9-year study period (collected date of CRS onset over 9 years).
Analytical Methods	Analytical methods appropriate for study design	-Yes. -Described model-building procedure. -Analytical methods accounted for sampling weights.	
	Report measures of precision and variability	-Presented variability of characteristics by case status. -95% Confidence intervals for odds ratios.	
	Report which statistical tests were used	-Odds ratios for multilevel logistic regression. -$p<0.05$ for significance testing.	
	Perform analysis to test sensitivity of results to alternative specifications.	-Assessed potential for residual confounding using negative outcome control. -Tested multiple averaging periods. -Tested impact of excluding participants with past symptoms.	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Discussion adequately addresses study limitations	Discussed potential for spatial confounding, lack of published work supporting conclusions, inability to account for conventional oil and gas operations, potential residual confounding, and potential for information bias.	No discussion of potential confounding by other potential environmental sources.
	Appropriate and complete interpretation of results	-Interpreted reported effect estimates appropriately.	-No discussion of other explanations for observed associations. -Interpretation does not consider multiple hypothesis testing.

MAGUIRE AND WINTERS 2017

Research Objective. The objective of this study was to assess the association between unconventional and conventional oil and natural gas development on self-reported life-satisfaction and bad mental health days.

Study Period and Location. The study took place in Texas with a separate examination of the Dallas-Fort Worth metropolitan area during the 2005–2010 time period.

Study Population. The study population consisted of Behavioral Risk Factor Surveillance Survey (BRFSS) respondents (the non-institutionalized population 18–85 years of age) living in the 153 of 254 Texas counties with large respondent samples (to preserve confidentiality) between 2005 and 2010. The survey represents a cross-section of the population, which differs for each survey year.

Outcome Ascertainment. Study participants provided: (1) life-satisfaction on a Likert scale ranging from one to four and, (2) bad mental health days in the past month.

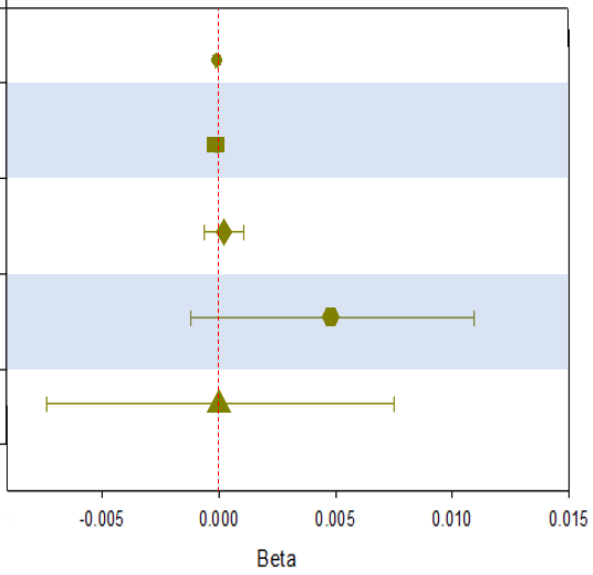
Exposure Assessment. The investigators used Texas Railroad Commission drilling permit data and the extracted spud date for each conventional and unconventional well, then aggregated this information to the month-county level. The primary exposure variables were count of drilled conventional wells and count of unconventional wells within the county of the respondents' residence in the last 12 months based on when a response was received. These two variables were analyzed separately. A second exposure variable, count of conventional or unconventional wells normalized by county area, was also assigned to each respondent.

Analytical Methods and Covariate Inclusion. The investigators used mixed linear regression, with a fixed effect for county and month-year. The model included the following covariates: sex, race/ethnicity, 5-year age group, marital status, educational attainment, number of adults in the household, and household child-adult ratio. A secondary analysis included income, employment status, county-level unemployment rate, and population density. Results are presented for the full state and for geographic subsamples, stratified by sex.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

Maguire and Winters et al. 2017 [1,2,3]
Betas and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Life Satisfaction

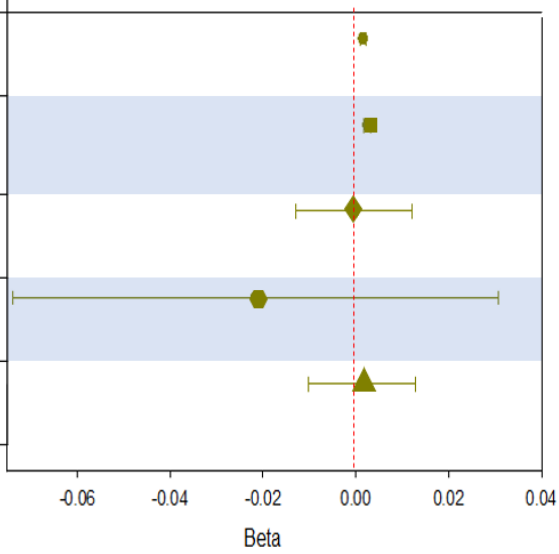
Model Feature	n
Full Sample	56,318
Dallas-Fort Worth metropolitan area	8,113
Full sample excluding Dallas-Fort Worth metropolitan area	48,205
Large Metropolitan Statistical Areas excluding Dallas-Fort Worth metropolitan area	18,861
Small and Medium Metropolitan Statistical Areas	29,344



1. Plotted from results as presented in each study (excluding supplemental information).
2. Each color represents an outcome, and each shape represents a model.
3. Exposure Surrogate: Horizontal well count within county of residence. (Total n = 56,318; Exposure Group: Continuous exposure).

Maguire and Winters et al. 2017 [1,2,3]
Betas and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Bad Mental Health Days

Model Feature	n
Full Sample	57,940
Dallas-Fort Worth metropolitan area	8,392
Full sample excluding Dallas-Fort Worth metropolitan area	49,548
Large Metropolitan Statistical Areas excluding Dallas-Fort Worth metropolitan area	19,453
Non-Metropolitan Statistical Areas; Small and Medium Metropolitan Statistical Areas	30,095



1. Plotted from results presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Horizontal well count within county of residence. (Total n = 57,940; Exposure Group: Continuous exposure).

Maguire and Winters 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	-Participants recruited using random sampling methods. -Included participants should be representative of Texas population living in counties with larger populations.	Unclear if investigators used sampling weights provided by BRFSS to decrease likelihood of selection bias and effectiveness of random sampling methods.
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on size of county-level population, age.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to cross-sectional study design.	
	Control group appropriate to address study question	Not applicable to cross-sectional study design.	
	Same population over study period		No formal assessment of this assumption for the 5-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)		Not reported.
Outcome Assessment	Outcome ascertained using valid and reliable measures	Depends on validity or reliability of BRFSS data.	-Not medically confirmed and possibly subject to information bias. -No discussion of validity or reliability of life-satisfaction or number of bad mental health dates scales.
	Outcome assessors blinded to exposure status	BRFSS is conducted without knowledge of this particular analysis.	
	No systematic differences in outcome ascertainment or reporting between exposure groups		Unclear whether there is differential participation in the BRFSS by exposure status or other factors.
Exposure Assessment	Performed using valid, reliable and sensitive methods	Dependent on the quality of underlying data from Texas Railroad Commission.	-Assumed residential stability throughout potentially exposed period. -No discussion of data quality.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Included all natural gas wells (objective to study all oil and gas wells). -Created separate surrogates	- No evaluation of whether exposure surrogate represents OGD activity.

Maguire and Winters 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
		for conventional and "horizontal" wells.	
	Study period sufficient to capture exposure variability	Not applicable to cross-sectional study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	Used continuous metrics.	
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Included only wells drilled in previous year.	May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		Unclear if one year is appropriate time-scale for "life-satisfaction" and "bad mental health days."
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Collected basic individual-level demographic, SES characteristics, household occupants using administered questionnaire.	
	Controlled for baseline conditions	Controlled for basic individual-level demographic and SES characteristics.	Lacking detailed information on co-morbidities, SES, family medical history, and lifestyle factors.
	Controlled for background exposures	Controlled for number of conventional wells drilled in prior year.	No control of other potential environmental sources, including occupational exposures, industrial sources, or traffic.
	Assessed time trends	Controlled for month and year in model.	No temporal trends described.
Analytical Methods	Analytical methods appropriate for study design	-Methods account for intra-county correlation. -Present rationale for chosen method.	Outcomes did not follow a normal distribution, and therefore linear model may not be appropriate.
	Report measures of precision and variability	95% Confidence intervals for odds ratios.	
	Report which statistical tests were used	-Beta coefficients for multilevel linear regression. -$p<0.01$ and 0.10 for significance testing.	
	Perform analysis to test sensitivity of results to alternative specifications.	-Stratified by geographic area and sex. -Tested using an ordered probit model. -Tested addition of employment, income and population density variables.	
Results and Discussion	All findings reported for analyses described in paper	All results provided for main analysis	Probit analysis results not presented.

Maguire and Winters 2017: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Discussion adequately addresses study limitations	Discussed potential for selection bias, residential mobility, limitations of exposure assessment	No discussion of potential for residual confounding
	Appropriate and complete interpretation of results	Interpreted reported effect estimates appropriately.	No discussion of alternative explanations for observed associations.

CASEY ET AL. 2018A

Research Objective. The objective of this study was to examine associations between Oklahoma earthquakes and statewide anxiety in Oklahoma measured by Google queries.

Study Period and Location. This study was conducted statewide in Oklahoma. Data were collected from January 2010 to May 2017.

Study Population. The study population included all individuals in Oklahoma who used the Google search engine for anxiety-related internet searches.

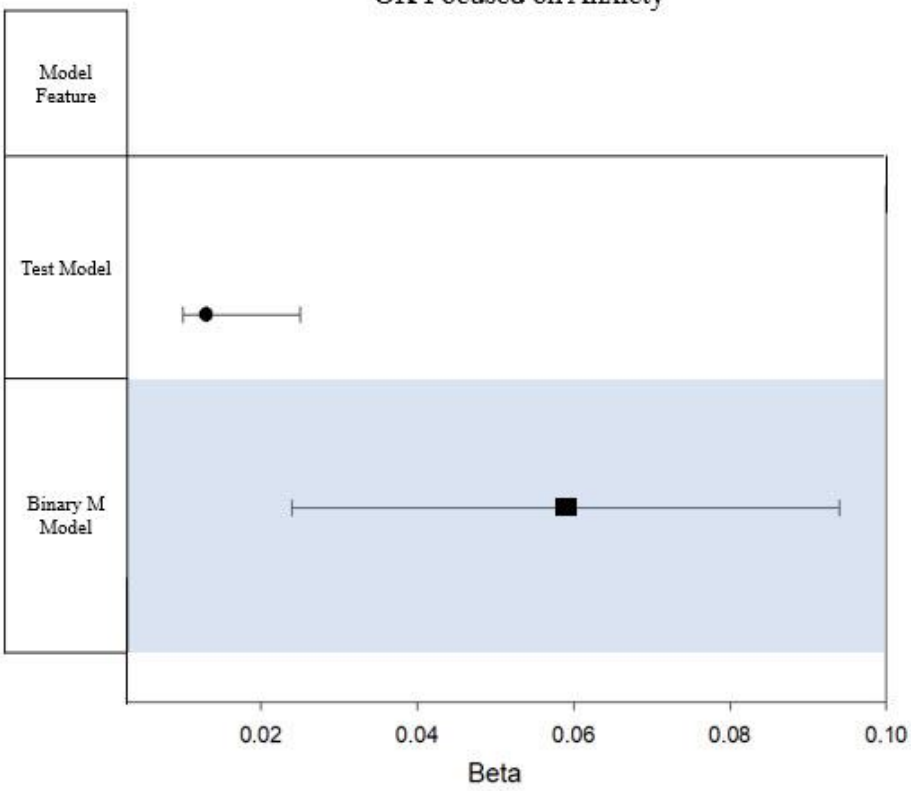
Outcome Ascertainment. The outcome was defined as Google searches for "anxious" or "anxiety" and a subsequent visit to a health-related website.

Exposure Assessment. The investigators assessed exposure as monthly earthquakes ≥4 in magnitude in the month of or before Google search queries. Number of earthquakes≥4 in magnitude was included as a binary and a continuous variable in analytical models.

Analytical Methods and Covariate Inclusion. The investigators used a time-series analysis approach, estimating monthly changes in the proportion of queries for anxiety. The investigators regressed monthly changes in the Oklahoma Google anxiety search episodes on two covariates (proportion of Google anxiety-related search episodes for the entire United States and within-Oklahoma monthly Google search episodes for "toothache"). Next the investigators used Box-Jenkins methods to identify and specify potential auto-correlation. The base model predicted changes in the monthly proportion of Google search episodes in Oklahoma. The investigators then included the monthly difference in earthquakes ≥4 in magnitude to the base model.

Results. The following figure summarizes results as presented by the study investigators (excluding any results provided in supplementary information).

Casey et al. 2018(a)[1,2,3]
Betas and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Monthly Proportion of Google Search Episodes in OK Focused on Anxiety



1. Plotted from results as presented in the study (excluding supplemental information).
2. Each shape represents a different model.
3. Exposure Surrogate: Number of USGS-recorded earthquakes ≥ 4 in magnitude within the state of Oklahoma. (Total n = 3.8 million).

Casey et al 2018a: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Includes Google searches from any Oklahoma location.	Unclear if those with computer access and those who use Google are representative of the general Oklahoma population.
	Inclusion/Exclusion criteria specified	Inclusions based on Google search phrase.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to ecologic study design.	
	Control group appropriate to address study question	Not applicable to ecologic study design.	
	Same population over study period		No formal assessment of this assumption for the 7-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)		Covariates not collected.
Outcome Assessment	Outcome ascertained using valid and reliable measures		-Google searches may represent an interest in anxiety and may not represent clinically significant anxiety, resulting in important outcome misclassification. -Validity of outcome ascertainment approach not assessed.
	Outcome assessors blinded to exposure status		Assumes people searching for anxiety would have felt earthquake
	No systematic differences in outcome ascertainment or reporting between exposure groups		Assumes that people searching for anxiety information are the same as those who felt earthquakes.
Exposure Assessment	Performed using valid, reliable and sensitive methods	Dependent on the quality of underlying USGS Advanced National Seismic System's data.	Assumed residential stability throughout study period.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	Include only magnitude ≥4 earthquakes	Unclear if all earthquakes are associated with UOGD.
	Study period sufficient to capture exposure variability	Study conducted over a period of UOGD variability.	
	Selection of exposure groups that represent the	Choice of magnitude ≥4 earthquakes sufficient to detect earthquake.	Unclear if those searching for "anxiety" would have felt an

Casey et al 2018a: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	full range of variability in UOGD.		earthquake between 3 and 4 magnitude.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells		No.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		No data collected on when symptoms appeared for those searching for anxiety information.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Included searches for "toothache."	No others assessed.
	Controlled for baseline conditions		No,
	Controlled for background exposures		No control of other potential environmental sources, including occupational exposures, industrial sources, noise or light pollution, or conventional wells.
	Assessed time trends	-Included Google anxiety-related search episodes for the United States to test for trends. -Included Google search episodes for "earthquakes" to control for interest in earthquakes over time.	
Analytical Methods	Analytical methods appropriate for study design	-Yes. -Analytical methods appropriate for assessment of time-series data. -Transformed variables to fit assumptions of time-series model.	
	Report measures of precision and variability	-Presented variability of earthquakes over time. -95% Confidence intervals for beta coefficients.	
	Report which statistical tests were used	Yes.	

Casey et al 2018a: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Perform analysis to test sensitivity of results to alternative specifications.	-Assessed model omitting OK "toothache" searches. -Checked for influence of outliers in dependent variable. -Included a negative exposure variable (earthquakes <=2.5 magnitude) to test for residual confounding. -Included alternative exposure specifications.	
Results and Discussion	All findings reported for analyses described in paper		Evidence that time-series variable were appropriately transformed fit model assumptions were not provided.
	Discussion adequately addresses study limitations	Discussed inability to detect clinical mental health outcomes, potential outcome misclassification, and residual confounding.	-No discussion of validity of using Google searches to ascertain the outcome, potential for selection bias, exposure misclassification (i.e., relationship between proximity to earthquake epicenter and propensity to perform a Google search). -No discussion of limitations of ecologic-level assessment.
	Appropriate and complete interpretation of results	Interpreted reported effect estimates appropriately.	No discussion of rise in anxiety searches in 2017 with simultaneous decline in number of earthquakes in the same year.

CASEY ET AL. 2018B

Research Objective. The objective of this study was to evaluate the association between unconventional natural gas development with depression symptoms and disordered sleep diagnoses.

Study Period and Location. The investigators collected cross-sectional survey data in 2014-2015 from participants living in central and northeastern Pennsylvania. Additional information for survey respondents was retrieved from respondents' Electronic Medical Records for January 2009- June 2015.*Study Population.* The study population included patients in the Geisinger Health system. Of 23,700 letters sent to potential participants, 7,847 participants responded to the primary questionnaire and 4,966 participants responded to the secondary questionnaire. The final analytical sample included 4,932 participants residing in Pennsylvania. Racial/ethnic minorities and individuals with higher risk of chronic rhinosinusitis were oversampled

Outcome Ascertainment. The investigators evaluated two primary outcomes: depression symptoms (severity) and disordered sleep. Depression symptoms were evaluated using a patient health questionnaire (PHQ-8). Disordered sleep diagnoses were identified in Geisinger's Electronic Health Records based on ICD-9 codes for disordered sleep and orders for medication classified as "hypnotics."

Exposure Assessment. The investigators collected well location, dates of well pad construction, drilling, stimulation, well depth, and volume of natural gas production from Pennsylvania Department of Environmental Protection, Department of Conservation and Natural Resources and SkyTruth. Investigators created an IDW-squared surrogate for four phases of well development and summed it for the 14 days and 90 days prior to the survey response and sleep diagnosis, respectively. Investigators created a summary *z*-score for each of the four phases and modeled the surrogate as a categorical variable in quartiles.

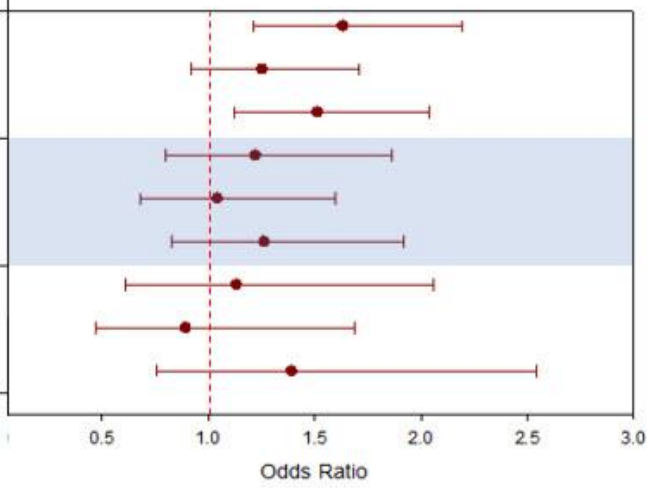
Analytical Methods and Covariate Inclusion. The investigators used multinomial logistic models to estimate the association between the surrogate measures of exposure and depression symptoms. Investigators also used a negative binomial model to evaluate depression symptoms continuously. The investigators used a survey-weighted generalized estimating equations model for the disordered sleep analysis. Sensitivity analyses assessed the influence of sample weights. Investigators included the following covariates in models: race/ethnicity, sex, Medical Assistance (i.e., Medicaid), age, smoking status, alcohol use status, body mass index, community socioeconomic deprivation, and water source.

Results. The following figures summarize results as presented by the study investigators (excluding any results provided in supplementary information).

Casey et al. 2018(b)[1,2]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Depression Symptom Severity

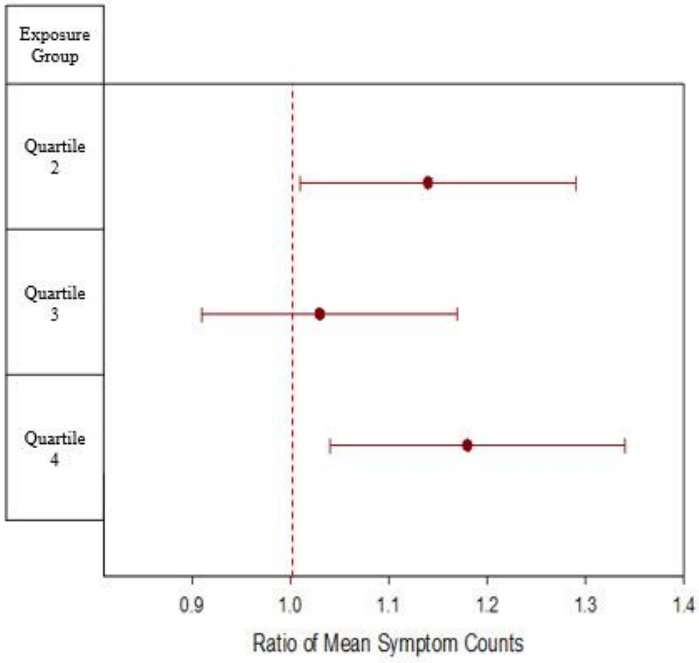
End Point	Cases	Exposure Group
Mild Depression Symptoms	1,075	Quartile 2
		Quartile 3
		Quartile 4
Moderate Depression Symptoms	454	Quartile 2
		Quartile 3
		Quartile 4
Moderately Severe/Severe Depression Symptoms	257	Quartile 2
		Quartile 3
		Quartile 4



1. Plotted from results as presented in the study (excluding supplemental information).
2. Exposure Surrogate: Inverse distance weighted Z-score transformed activity phase. (Referent: Quartile 1; Total n = 4,762).

Casey et al. 2018(b)[1,2]

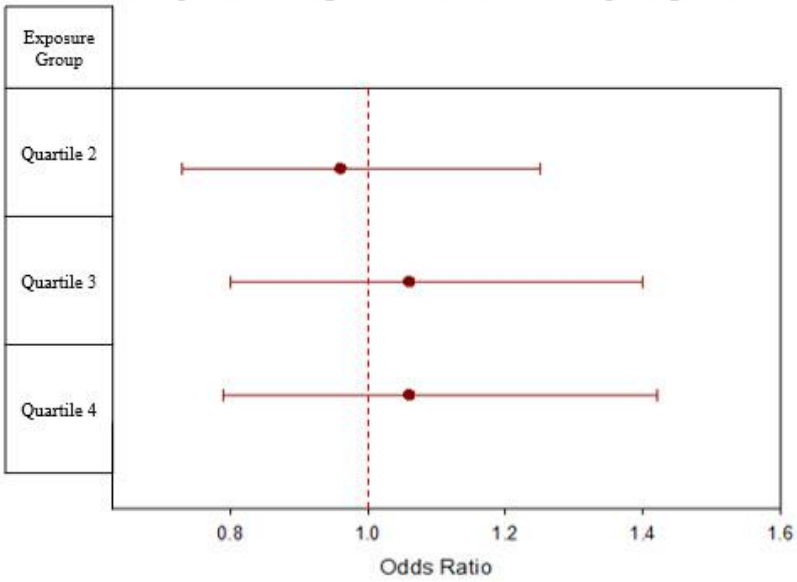
Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Depression Symptom Count



1. Plotted from results as presented in the study (excluding supplemental information).
2. Exposure Surrogate: Inverse distance weighted activity phase. (Referent: Quartile 1; Total n = 4,762).

Casey et al. 2018(b)[1,2]

Odds Ratios and 95% Confidence Intervals for the Relationship between Exposure Surrogate and Disordered Sleep Diagnosis



1. Plotted from results as presented in the study (excluding supplemental information).
2. Exposure Surrogate: Inverse distance weighted Z-score transformed activity phase. (Referent: Quartile 1; Total n = 3,658).

Casey et al 2018b: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Study population obtained from Geisinger Health System.	-Unclear if survey responders are representative of general PA population. -Oversampled people with nasal and sinus symptoms and racial/ethnic minorities. -Participation may depend on symptom severity (selection bias).
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Excluded respondents with residence outside of study area and those who did not answer depression symptom questions.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to cross-sectional study design.	
	Control group appropriate to address study question	Not applicable to cross-sectional study design.	
	Same population over study period		No formal assessment of this assumption over 6-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: race and BMI	-Percentage female, Medical Assistance, current smoker status, residence in a city, depression medication use higher in depression symptoms groups. -Percentage well water use and moderate alcohol use highest in no depression symptoms group.
Outcome Assessment	Outcome ascertained using valid and reliable measures	-PHQ-8 is a validated questionnaire to measure of depression symptoms. -Ascertained sleep disorder outcomes using ICD-9 diagnosis coding and medication use.	-Depression symptoms ascertained via self-report, which may be subject to bias. -Medications may be used for ailments other than disordered sleep, resulting in potential outcome misclassification.
	Outcome assessors blinded to exposure status	-Disordered sleep diagnoses and medication use recorded separately from exposure assignment. -Study questionnaire did not mention UNGD in explanation of study objectives.	
	No systematic differences in outcome ascertainment or reporting between exposure groups		Not specified.

Casey et al 2018b: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Exposure Assessment	Performed using valid, reliable and sensitive methods	-Dependent on the quality of underlying data from Pennsylvania Department of Environmental Protection and Pennsylvania Department of Conservation and Natural Resources. -Used crowdsourcing to identify well pad location. -Geocoded to home address.	No discussion of quality of data.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No
	Assess exposure in a way that addresses review question.	-Incorporated information about proximity to and number of wells. -Included only UNGD wells. -Assigned at daily-resolution.	No evaluation of whether exposure surrogate represents UNGD activity.
	Study period sufficient to capture exposure variability	Not applicable to cross-sectional study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.		-Limited spatial overlap between study population and UNGD wells. -Did not test for cut-point bias.
	Differentiates among UOGD and its various phases	Reflected pad preparation, spud, stimulation, and production phases.	-Pad development and phase duration estimated. - Unable to distinguish between phases because of collinearity (aggregated phases to z-score).
	Differentiates between active and non-active wells		Not reported.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		Unclear if 14 days and 90 days are sufficient to develop depressive symptoms and disordered sleep, respectively.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Used electronic health record data and administered questionnaires to collect individual-level covariate information, and Census and PADEP data to collect community-level data.	
	Controlled for baseline conditions	-Collected use of anti-depressant medications. -Controlled for basic individual- and community-level SES, individual-level demographic	Lacking detailed information on co-morbidities, SES, family medical history, and lifestyle factors (except smoking and alcohol).

Casey et al 2018b: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
		characteristics, and alcohol consumption.	
	Controlled for background exposures	Controlled for water supply source and smoking.	No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells.
	Assessed time trends	Restricted analysis to 2014-2015.	No assessment of change in prevalence of covariates over time.
Analytical Methods	Analytical methods appropriate for study design	-Yes. -Described model-building procedure. -Analytical methods accounted intra-individual correlation and sampling weights.	
	Report measures of precision and variability	-Presented variability of characteristics by exposure group. -95% Confidence intervals for odds ratios.	Temporal variability of exposure not presented.
	Report which statistical tests were used	-Odds ratios for multilevel logistic regression. -$p<0.05$ for significance testing.	
	Perform analysis to test sensitivity of results to alternative specifications.	-Assessed effect modification by antidepressant medication use. -Tested impact of sampling weights. -Excluded residence with imprecisely geocoded addresses.	
Results and Discussion	All findings reported for analyses described in paper	Yes.	
	Discussion adequately addresses study limitations	Discussed potential for selection bias, residential mobility, and limitations of exposure assessment, unequal distribution of covariates between respondents and non-respondents, lack of data on perceptions of UOGD or whether respondents were leaseholders.	-No discussion of hydrology or meteorology, potential influence of other environmental sources on results.
	Appropriate and complete interpretation of results	-Interpreted reported effect estimates appropriately. -Discussed some external factors that may explain study results.	

ELLIOTT ET AL. 2018

Research Objective. The primary study objective was to examine associations between residential proximity to unconventional oil and natural gas wells and concentrations of analytes measured in Ohio well water. As a secondary objective, the investigators evaluated associations between residential unconventional oil and gas proximity and health symptoms.

Study Period and Location. This study was conducted in Belmont County, Ohio. Data were collected in 2016.

Study Population. The study population included 66 residents of Belmont County, Ohio, recruited using mailed informational flyers, local newspaper and television news stories, and social media. The following inclusion criteria were required for participation: ≥ 21 years of age, head of household, and English-speaking. The investigators preferentially enrolled participants with groundwater as their drinking water source.

Outcome Ascertainment. Trained study interviewers administered a questionnaire to collect information on the following symptom groups: respiratory (e.g., allergies and wheezing), dermal (e.g., skin rash ≥ 3 days and burning skin), neurologic (e.g., severe headaches and dizziness), gastro-intestinal (e.g., stomach ulcers and nausea), and general (e.g., stress and fatigue).

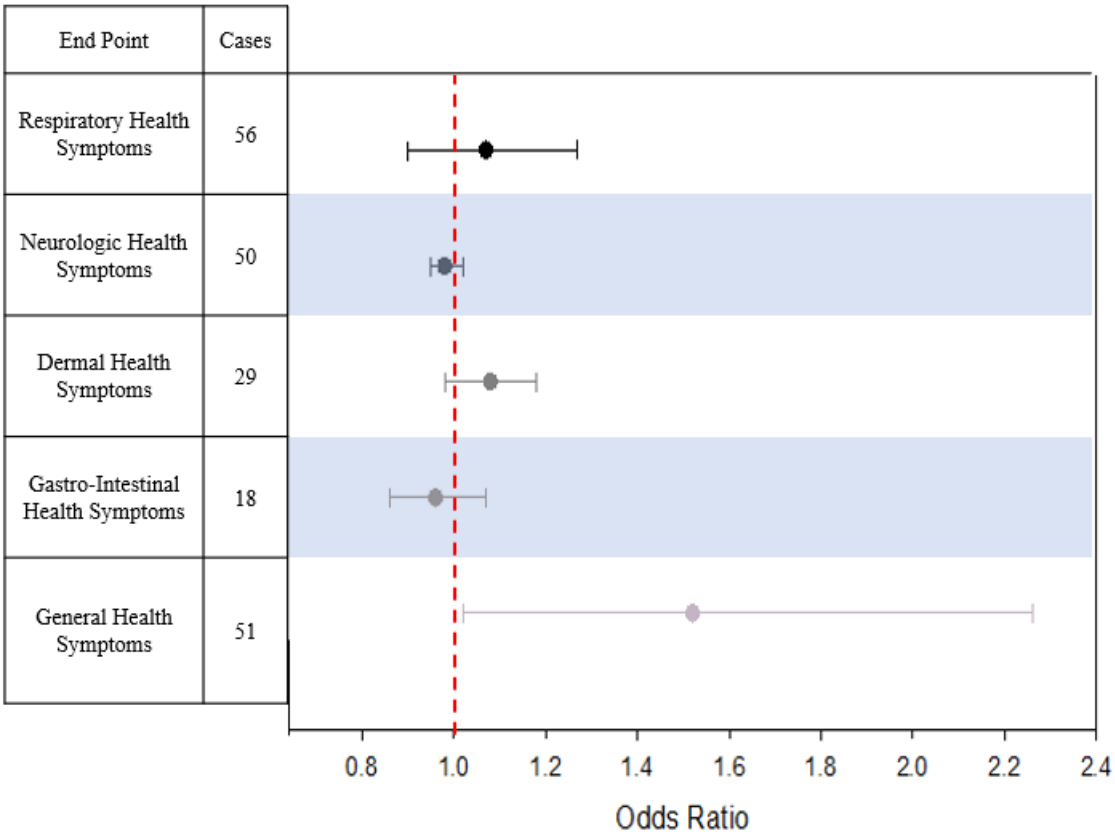
Exposure Assessment. The investigators used data from the Ohio Department of Natural Resources to construct three exposure surrogates: distance (km) to nearest active unconventional oil and gas well, inverse-distance weighted well count, and inverse-distance-squared weighted well count within a 5-km radius of the participant residential address. The investigators calculated metrics specific to the drilling/drilled or production phases and explored all inverse-distance weighted metrics, along with alternative radii of 1 km and 2 km.

Analytical Methods and Covariate Inclusion. The investigators fit multivariable logistic regression models for all exposure surrogate and health outcomes. They considered the following covariates for model inclusion, using backwards selection for model building: age, sex, body mass index, smoking status, educational status, marital status, and employment status. Sensitivity analyses explored whether there were any differences in associations using phase-specific metrics (drilling/drilled or production).

Results. The following figure summarizes results as presented by the study investigators (excluding any results provided in supplementary information).

Elliott et al. 2018[1,2,3]

Odds Ratios and 95% Confidence Intervals for the Relationship Between Exposure Surrogate and Self-Report Health Symptoms



1. Plotted from results presented in the study (excluding supplemental information).
2. Each color represents a different outcome.
3. Exposure Surrogate: Inverse distance weighted-squared within Belmont, County Ohio. (Total n = 66).

Elliott et al 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Participants recruited using random sampling methods	-Unclear what population investigators aim to study. -Subject to selection bias: average age of the study sample over 60 years
	Inclusion/Exclusion criteria specified	-Detailed discussion of the criteria used to select study sample. -Exclusions based on age, language spoken at home, residential location (outside of Belmont County), and water source.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to cross-sectional study design.	
	Control group appropriate to address study question	Not applicable to cross-sectional study design.	
	Same population over study period	One-time survey, with all surveys conducted over the course of 2-3 months.	
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)		Population characteristics not presented by exposure groups.
Outcome Assessment	Outcome ascertained using valid and reliable measures		-Not medically confirmed and possibly subject to information bias. -Unclear which specific outcomes investigators measured.
	Outcome assessors blinded to exposure status		Unclear to what extent the participants knew about the study objectives with regards to UOGD exposures
	No systematic differences in outcome ascertainment or reporting between exposure groups		Not specified.
Exposure Assessment	Performed using valid, reliable and sensitive methods	-Dependent on the quality of underlying data Ohio Department of Natural Resources. -Geocoded to home address.	-Assumed residential stability throughout exposure period. -No discussion of quality of data.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Examined correlations between exposure surrogate and concentrations of	Did not use water samples in health analysis.

Elliott et al 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
		analytes collected in tap water samples. -Incorporated information about proximity to and number of wells. -Included only UOGD wells.	
	Study period sufficient to capture exposure variability	Not applicable to cross-sectional study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.	-Used continuous metrics. -Tested various distances from the residence to calculate the exposure surrogate.	
	Differentiates among UOGD and its various phases	Considered drilling and production phases in sensitivity analysis.	-Limited explanation of methods used to determine drilling and production phases. -Missing information about pad preparation and stimulation.
	Differentiates between active and non-active wells	Included only active UOGD wells study area	May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed	Investigators did not consider outcomes with long latency periods.	No temporal component between exposure and effect considered.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Collected a wide range of current demographic, SES, and lifestyle factors using administered questionnaire.	
	Controlled for baseline conditions	Controlled for one or two variables for most health symptoms category.	-Family or health history not collected. -Limited control of any covariates. -No control of baseline conditions for respiratory or neurological symptoms.
	Controlled for background exposures		No control of other potential environmental sources, including occupational exposures, industrial sources, traffic, or conventional wells.
	Assessed time trends	Restricted analysis to one year.	No.
Analytical Methods	Analytical methods appropriate for study design	-Logistic regression appropriate for study design -Described model-building procedure.	
	Report measures of precision and variability	-95% Confidence intervals for odds ratios.	Temporal variability of exposure or covariates not presented.
	Report which statistical tests were used	-Odds ratios for logistic regression.	

Elliott et al 2018: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
		-$p<0.05$ for significance testing.	
	Perform analysis to test sensitivity of results to alternative specifications.	Stratified by sampled water source.	
Results and Discussion	All findings reported for analyses described in paper	All results provided for main analysis	-Results are not presented for stratified analyses. -Unclear whether all health symptom results were reported.
	Discussion adequately addresses study limitations	Discussed potential for selection bias, small sample size, multiple comparisons.	No discussion of hydrology, limited control of confounding, limitations of exposure assessment, whether respondents were leaseholders.
	Appropriate and complete interpretation of results	Interpreted reported effect estimates appropriately.	-No discussion of alternative explanations for observed associations. -Limited discussion of results from exposure surrogates and water analyte concentration models.

JEMIELITA ET AL. 2015

Research Objective. The objective of this study was to assess the association between UOGD and hospitalization prevalence rates.

Study Period and Location. Investigators collected ZIP-code level health and well location data between 2007 and 2011 for three Pennsylvania counties: Bradford, Susquehanna, and Wayne.

Study Population. The study population included all residents of ZIP codes located in the three counties (n = 67).

Outcome Ascertainment. The investigators obtained inpatient hospital discharge records (n = 92,805) and ICD-9 coding from the Pennsylvania Healthcare Cost Containment Council. Patient records were excluded for the following diagnostic related group codes: dentistry, human immunodeficiency virus (HIV), and neurosurgery. Inpatient prevalence rates by ZIP code were calculated by dividing the ICD-9-specific inpatient counts per year by the ZIP-code population in that year (obtained from census data).

Exposure Assessment. The investigators identified gas wells defined as unconventional from the Pennsylvania Department of Environmental (PADEP) Protection Oil and Gas Reporting website. Wells were mapped, and the number of active wells and well density per ZIP code were calculated for each year between 2007 and 2011.

Analytical Methods and Covariate Inclusion. For both well count and well density analyses, the investigators used conditional fixed effects Poisson regression, with ZIP codes as the fixed effects. Well count was included as a linear predictor and as a quadratic predictor in separate models. Well density was included as a categorical predictor (by quartiles). The investigators presented risk ratios and P values, with a Bonferroni correction of $P < 0.00096$. Year of the inpatient record was included as a covariate in all models. The investigators did not report measures of precision.

Results. The investigators did not present results with measures of uncertainty (e.g., confidence intervals); therefore, they cannot be presented in a summary plot.

Jemielita et al. 2015: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Study Population	Study population representative of underlying population	Included all hospitalization records from study area.	
	Inclusion/Exclusion criteria specified	-Inclusion and exclusion criteria specified. -Exclusions based on residence outside of study area, and records with dentistry, HIV, or neurosurgery diagnoses.	
	Attrition not systematically different between exposure groups (cohort studies) or cases and controls (case-control studies)	Not applicable to ecologic study design.	
	Control group appropriate to address study question	Not applicable to ecologic study design.	
	Same population over study period		No assessment of this assumption over four-year study period.
	Baseline characteristics similar between exposure groups (cohort studies) or cases and controls (case-control studies)	Similar: county-level age, educational attainment, median income, proportion male averaged across study period.	Percentage Black higher in the county without UOGD compared to two counties with UOGD.
Outcome Assessment	Outcome ascertained using valid and reliable measures	-Ascertained using ICD-9 and Medicare Severity Diagnosis Related Groups (MS-DRGs) codes from inpatient discharge records. -Discussed quality of data source.	-Unclear how investigators determined medical categories and combined ICD-9 and MS-DRG codes that overlapped. -No discussion of quality of MS-DRG data.
	Outcome assessors blinded to exposure status	Ascertained without knowledge of exposure status.	
	No systematic differences in outcome ascertainment or reporting between exposure groups	Collected separately from exposure.	
Exposure Assessment	Performed using valid, reliable and sensitive methods	Dependent on quality of underlying data (PADEP)	-Assumed residential stability throughout study period. -No discussion of quality of PADEP data.
	Non-differential between outcome groups	Yes.	
	Includes measurements of chemical and non-chemical agents		No.
	Assess exposure in a way that addresses review question.	-Provided clear definition of unconventional wells. -Included information about well count and density.	No individual-level data.

Jemielita et al. 2015: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
	Study period sufficient to capture exposure variability	Not applicable to ecologic study design.	
	Selection of exposure groups that represent the full range of variability in UOGD.		Did not test for cut point bias for density analysis.
	Differentiates among UOGD and its various phases		No.
	Differentiates between active and non-active wells	Yes.	May account for wells that become inactive later in study period.
	Timeframe sufficient to expect to see an association between exposure and outcome if it existed		No consideration of timing of exposure in relation to outcomes assessed for 25 different medical categories, which may have different relevant times-at-risk.
Confounding	Potential confounding variables assessed comprehensively and consistently across exposure groups (cohort studies) or cases and controls (case-control studies)	Collected ZIP-code level basic demographic characteristics and median income.	No control of any potential confounders analytically.
	Controlled for baseline conditions		No control of ZIP-code level baseline conditions.
	Controlled for background exposures		No control of other potential environmental sources, including traffic, conventional wells, industrial sources or other individual co-exposures.
	Assessed time trends		Not assessed.
Analytical Methods	Analytical methods appropriate for study design	Yes: Fixed effects Poison regression with robust standard errors (controls for non-time-varying ZIP-code level characteristics).	No potential confounders included in model.
	Report measures of precision and variability	-Presented variability in inpatient prevalence rates by medical category. -Provided visual display of inpatient prevalence rates by ZIP-code.	No measures of uncertainty provided.
	Report which statistical tests were used	-Wald test -Hypothesis testing at $p<0.05$ -Bonferroni correction for multiple comparisons.	
	Perform analysis to test sensitivity of results to alternative specifications.	Tested removal of outlier ZIP-codes.	

Jemielita et al. 2015: Important Strengths and Limitations Noted by the Committee			
Category	**Criteria**	**Strengths**	**Limitations**
Results and Discussion	All findings reported for analyses described in paper		-Unclear which outcomes (i.e., ICD-9 codes) were omitted from analysis. -Do not present analysis with wells included as a quadratic term.
	Discussion adequately addresses study limitations	Mentioned population mobility, potential exposure misclassification, prevalence rate trends over time.	- -Analysis not targeted to outcomes relevant to study objectives. -Did not discuss of limitations of ecologic study design.
	Appropriate and complete interpretation of results	Interpreted reported effect estimates appropriately	-Inappropriate inferences given small magnitude of results and study limitations. -Interpretation does not consider multiple hypothesis testing.