ATTACHMENT 32

Final Report

for

Pennsylvania Department of Health,

Bureau of Epidemiology

Hydraulic Fracturing Epidemiology Research Studies: Asthma Outcomes

Prepared by:

University of Pittsburgh

School of Public Health

Contract number: 4400018535

July 31, 2023

List of Contributors

Faculty
Jeanine M. Buchanich, MPH, MEd, PhD, Associate Professor, Department of Biostatistics
Principal Investigator, Retrospective Cohort Studies

Evelyn O. Talbott, MPH, DrPH, Professor of Epidemiology
Principal Investigator, Childhood Cancer Case-Control Study

Vincent Arena, PhD, Associate Professor Emeritus, Department of Biostatistics

Todd M. Bear, PhD, MPH, Assistant Professor, Department of Family Medicine, School of Medicine

James P. Fabisiak, PhD, Associate Professor, Department of Environmental and Occupational Health

Sally E. Wenzel, MD, Chair, Department of Environmental and Occupational Health

Ada O. Youk, PhD, Associate Professor, Department of Biostatistics

Jian-Min Yuan, MD, PhD, Professor, Department of Epidemiology

Staff
Jennifer Fedor, MS, Data Analysis

Allison C. Koller, MS, Project Coordinator

Michael F. Lann, MSIS, Programmer

David Maynard, Administrative Assistant

Andrew Mrkva, MA, Web Developer

Melanie Stangl, Administrative Assistant

Students
Jacob Cherry (Department of Biostatistics)
Violet Hecmanczuk (Department of Biostatistics)
Maggie Kuzemchak, MS (Department of Biostatistics)
Tianyuzhou (Jenny) Liang (Department of Biostatistics)
Emily Nicholls (Department of Environmental and Occupational Health)
Brian O'Connell, MS (Department of Biostatistics)
Nina Oryshkewych, MS (Department of Biostatistics)
Natalie Price (Department of Epidemiology)
Nicholas Tedesco (Department of Biostatistics)
Alex Watts (Department of Biostatistics)
Fan Wu, MPH (Department of Environmental and Occupational Health)

Consultant: Ravi Sharma, PhD, Adjunct Assistant Professor, Department of Epidemiology

Acknowledgments

The study team wishes to acknowledge and thank our External Advisory Board members for their thoughtful input throughout the study, including at meetings on August 5, 2021, November 10, 2021, January 26, 2022, April 13, 2022, and August 11, 2022.

External Advisory Board Members

David Allard, MS, Pennsylvania Department of Environmental Protection, Bureau of Radiation Protection (retired)

Uni Blake, PhD, American Petroleum Institute

Laura Dagley, Physicians for Social Responsibility (resigned)

Erica Jackson, FracTracker Alliance (resigned)

Ned Ketyer, MD, Physicians for Social Responsibility (resigned)

Jan Maund, (retired) Trinity Area School District

Curtis Schreffler, Deputy Director, Hydrologic Studies, PA Division, United States Geological Survey

Heaven Sensky, Center for Coalfield Justice (resigned)

Logan Spector, PhD, University of Minnesota Medical School

Judy Wendt Hess, PhD, Shell Oil

Edward Yorke, Canonsburg Councilman

We wish to thank the following organizations for providing study data. All interpretations are our own.

Pennsylvania Department of Health

Pennsylvania Department of Environmental Protection

University of Pittsburgh Biomedical Informatics Health Record Research Request (R3)

We also wish to thank Allison Hydzik, Media Relations, University of Pittsburgh Medical Center, for her invaluable assistance and advice.

Table of Contents

Tables

Figures

Asthma Outcomes Cohort Study

Background

Over the last 25 years, the American energy landscape has undergone an evolution, perhaps most notably with the expansion of hydraulic fracturing operations[1]. From 2000 to 2015, the number of hydraulically fractured wells in the United States increased from 23,000 to approximately 300,000. This rapid growth has corresponded to a range of economic benefits, including decreased energy costs and greatly increased production of both oil and natural gas[2]. However, mounting evidence suggests that hydraulic fracturing may have adverse impacts on public health and the environment[3-24].

Hydraulic fracturing – also known as fracking – is a process of unconventional natural gas development (UNGD) done by injecting large amounts of fluid at high pressure into dense rock in order to free trapped oil and natural gas[25]. The fluid used for injection typically consists of a mixture of water, sand (or other proppants), and various chemical additives. These wells, which are typically deeper than conventional wells, access previously unavailable reservoirs of oil and natural gas trapped in shale. The Marcellus Shale formation encompasses approximately half of Pennsylvania and is a large reservoir of natural gas.

Exposure to UNGD has been shown to be associated with some asthma exacerbations, including in Pennsylvania (PA)[8,13,26,27]. Rasmussen et al.[8] performed a case-control study using electronic health record data on 35,508 patients with asthma, aged 5 to 90 years, in Eastern PA. Patients with exacerbations from 2005 to 2012 were frequency matched on age, sex, and year of event to patients without an event. They assessed exposure to UNGD activity by well phase (well pad preparation, drilling, hydraulic fracturing, and production) the day before the event. Rasmussen et al. found statistically significant elevations in the highest tertile of activity, compared to the baseline of very low activity, for 11 of 12 UNGD phases by outcome models examined. The highest tertile odds ratios ranged from 1.45 for hospitalizations in the well pad preparation phase to 4.43 for mild exacerbations in the production phase.

Koehler et al.[13] (2018) used a principal components analysis to evaluate the association between mild asthma exacerbations, defined as new oral corticosteroid medications, and three UNGD activity metrics. They included nearly 70,000 exacerbations from approximately 40 counties in Eastern Pennsylvania. They constructed an exposure measure which included well pad development, drilling, stimulation, production, and compressor engines. Koehler et al. found statistically significantly elevated risk among those living within 1 kilometer (km) of the nearest well drilled, those in the highest tertile using an inverse distance weighted (IDW) cumulative metric, and those in the highest quartiles of exposure in the metrics that considered well phases plus distance.

There were three specific aims of this retrospective cohort study of asthma: 1) to replicate earlier studies conducted in Eastern PA using a population in Southwestern PA, where UNGD has proliferated in the past 15 years; 2) to enhance and improve upon previous UNGD exposure characterizations by assessing the associations between asthma exacerbations of various severity and each the four phases of UNGD; and 3) to enhance and improve upon previous UNGD exposure characterizations by assessing whether associations varied by multiple buffer distances to individuals' residences.

Methods

Asthma Records Data

Cohort members were identified from a University of Pittsburgh Health Record Research Request (R3) data request following University of Pittsburgh Institutional Review Board (IRB) approval. R3 is a service of the Department of Biomedical Informatics (DBMI) managed by the Chief Research Informatics Officer (CRIO), sponsored in part by the Clinical and Translational Sciences Institute and Institute for Precision Medicine at the University of Pittsburgh.

To be part of the cohort, participants need to have:

- An electronic health record with the University of Pittsburgh Medical Center (UPMC) health system between 2011-2020
- Age 5-90 years
- Patient residence within a zip code located within the eight-county study area (Allegheny excluding the City of Pittsburgh (excluded zip codes listed in Appendix Table 1), Armstrong, Beaver, Butler, Fayette, Greene, Washington, and Westmoreland counties)
- Primary diagnosis of asthma (codes shown in Appendix Table 2)
- At least one order for medications prescribed for asthma (Appendix Table 3)

We excluded participants with:

- Cystic fibrosis
- Pulmonary hypertension and pulmonary vascular disease (including pulmonary embolism)
- Paralysis of vocal cords or larynx
- Bronchiectasis
- Pneumoconiosis

ICD-9 and ICD-10 codes for inclusion and exclusion criteria are provided in Appendix Table 2.

Outcome Measures

Of interest were three levels of severity of asthma exacerbations among patients with asthma, defined according to the American Thoracic Society (ATS) and European Respiratory Society (ERS)[28]. Under the ATS/ERS criteria, patients with asthma are defined as those patients with a primary or secondary diagnosis of asthma (ICD-9 or ICD-10 codes; Appendix Table 2) in their electronic health record. Only exacerbation events among patients with at least one primary diagnosis were eligible for inclusion in the analysis. Exacerbation events were defined as follows:

1. **Severe exacerbation:** Initiation or increase of systemic corticosteroid medications among patients with asthma (Appendix Table 3).

2. **Emergency Department (ED) severe exacerbation:** ED or urgent care encounters for asthma that involve treatment with systemic corticosteroids among patients with asthma.

3. **Hospital exacerbation:** Hospitalizations for asthma that involve treatment with systemic corticosteroids among patients with asthma.

For patients with more than one type of exacerbation within 1 week, only the most severe exacerbation was retained. For patients with more than one exacerbation of a given type within a calendar year, one exacerbation of that type was randomly selected.

Control Selection

Controls were selected from patients in the study population. Patients with asthma who did not have an exacerbation during the study period were eligible to be controls for an exacerbation of any type. Patients with asthma who did have an exacerbation during the study period were eligible to be controls for a less severe exacerbation or an exacerbation of equal or greater severity up until the calendar year of their exacerbation.

Among eligible control patients, control events were a randomly selected contact date per calendar year per patient to replicate the methodology used in Rasmussen et al[8]. Contact dates were identified as all encounters with the health system recorded in the electronic health record (e.g., office visits, medication orders, procedures, tests, etc.).

Controls were frequency matched to cases by the following criteria: age category (5-12, 13-18, 19, 44, 45-61, 62-74, 75-90); sex (male, female); year of encounter.

Events

We restricted the pool of candidate case events to those among patients aged 5-90 years and living in a study area residence on the day of the event and the day prior. We randomly selected one residence for events associated in time with multiple residences (n=370). Finally, we randomly selected one event per type, per year, per patient to represent our final set of case events.

Control encounters were frequency matched to case events on patient age group, patient sex, and encounter year. We used 1:1 control: case matching for severe events, 2:1 matching for ED severe events, and 4:1 matching for hospitalization severe events.

Covariate Definitions

Clinical and demographic features of the patient and of the environment surrounding the patient's residence were included as covariates to control for potential confounding. Patient residences were extracted by R3 from the electronic health records and geocoded. Addresses for residences in rural zip codes were masked (a small amount of uncertainty was added by R3 to the latitude and longitude) prior to receipt of the data to avoid potential re-identification.

We received clinical and demographic covariates including patient sex, family history of asthma, and race/ethnicity. We also received clinical and demographic covariates that could change depending on the encounter, including age category, year of event, season of event, overweight and obesity status, smoking status, and Type II diabetes diagnosis. Event-level covariates included: year, season, age, BMI category, smoking status, average maximum temperature (degrees Celsius) recorded in the patient's county of residence on the day prior to the event, and community level socioeconomic deprivation index quartile. Covariate information is shown in Table 1.

Table 1. Covariates Included in the Analysis

Covariate	Definition
Patient level variables	
Patient sex	Male Female
History of asthma in patient's first-degree relatives (parents, siblings, offspring)	Yes No
Race (Self-reported race of the patient, categorized from 19 options)	White Black All other races Unknown
Ethnicity (Self-reported ethnicity of the patient)	Hispanic Not Hispanic Unknown
Variables that were dependent on the event date	
Event year	Calendar year in which the event occurred
Season in which the event occurred (based on month and day of the event)	Winter: December 22 – March 21 Spring: March 22 – June 21 Summer: June 22 – September 21 Fall: September 22 – December 21
Patient age category	Age in years at the time of the event, categorized as: 5-12, 13-18, 19-44, 45-61, 62-74, 75-90
Overweight and obesity status	Based on BMI calculated based on the weight in pounds and height in feet and inches at the event date or averaged across the visits before and after the event date[29] (Appendix Table 4).
Smoking status of the patient at the time of the event	Current Former Never Unknown
Type II diabetes diagnosis	Whether the patient had a diagnosis of type II diabetes (ICD-9 code 250.x0 and 250.x2 or ICD-10 code E11.x) at the time of the event (yes, no)
Maximum temperature on the previous day (°C)	Maximum recorded temperature in degrees Celsius on the date prior to the event date from the weather station nearest to each patient's residence. If data were missing for the nearest weather station, we used the county-level average maximum temperature.
Variables that were dependent on the event date and residence	
Community socioeconomic deprivation index	Quartiles (Q)1 – Q4 divided equally by the total number of communities in our study area

	Higher values of the index reflect greater community socioeconomic deprivation (Appendix Table 5 for details)

Exposure Measure

Unconventional natural gas development

The primary exposure measure was an inverse distance-weighted index of UNGD activity[6,8,11,13,15] up to 10 miles (or 16,093.4 m) of patient residence. Due to small numbers of asthma cases living within 0.5 miles of wells and the masking of rural geocodes performed by R3, we considered four buffer distances: 1 mile, 2 miles, 5 miles, and 10 miles in these models.

There are four phases of UNGD: well pad preparation, drilling, hydraulic fracturing, and production, which vary in duration and potential exposures. Information required to calculate the UNGD activity metric was obtained from the Pennsylvania Department of Environmental Protection (PA DEP) and Pennsylvania Department of Conservation and Natural Resources (PA DCNR).

1. **Well pad preparation -** the process of preparing a site where one or more wells were located. It is defined as the period 30 days before the first well on the pad is spudded.

2. **Drilling -** the creation of the wellbore. This phase begins on the well's spud date and ends on the drilling completion date.

3. **Hydraulic fracturing** (fracking, stimulation) **-** the process of injecting large volumes of water at high pressure into the wellbore to fracture the shale layer. This period is defined as beginning on the stimulation commencement date and ending on the stimulation completion date. Hydraulic fracturing may be repeated over time for a given well.

4. **Production -** the process of collecting natural gas or oil that, following hydraulic fracturing, travels through the wellbore to the surface. Production durations are variable; produced gas volume was represented as an average daily gas volume. A well was defined as being in production for reporting periods when production is indicated and reported production volume is non-zero.

Phase-specific UNGD metrics were calculated for each exacerbation using the following equations in Table 2.

Table 2. Definition of UNGD activity metric phase durations

Phase	Phase name	Calculation of phase-specific activity metric
1	Well pad preparation	Phase 1 metric for patient *j* event *k* = $\sum_{i=1}^{n} \frac{1}{d_{ijk}^2}$ Where: • *n* is the number of well pads in development within 10 miles of the residence of patient *j* on the day prior to event *k*

		• d^2_{ijk} is the squared distance (m^2) between well pad *i* and the residence of patient *j* at the time of event *k*
2	Drilling	Phase 2 metric for patient *j* event *k* = $\sum_{i=1}^{n} \frac{1}{d_{ijk}^2}$ Where: • *n* is the number of wells in the drilling phase within 10 miles of the residence of patient *j* on the day prior to event *k* • d^2_{ijk} is the squared distance (m^2) between well *i* and the residence of patient *j* at the time of event *k*
3	Hydraulic fracturing	Phase 3 metric for patient *j* event *k* = $\sum_{i=1}^{n} \frac{w_i}{d_{ijk}^2}$ Where: • *n* is the number of wells in the hydraulic fracturing phase within 10 miles of the residence of patient *j* on the day prior to event *k* • w_i is the depth (m) of well *i* • d^2_{ijk} is the squared distance (m^2) between well *i* and the residence of patient *j* at the time of event *k*
4	Production	Phase 4 metric for patient *j* event *k* =$\sum_{i=1}^{n} \frac{v_i}{d_{ijk}^2}$ Where: • *n* is the number of wells in production within 10 miles of the residence of patient *j* on the day prior to event *k* • v_i is the produced gas volume (m^3) of well *i* on the day prior to event *k* • d^2_{ijk} is the squared distance (m^2) between well *i* and the residence of patient *j* at the time of event *k*

Figure 1 illustrates the calculation of the phase-specific and buffer-specific metrics.

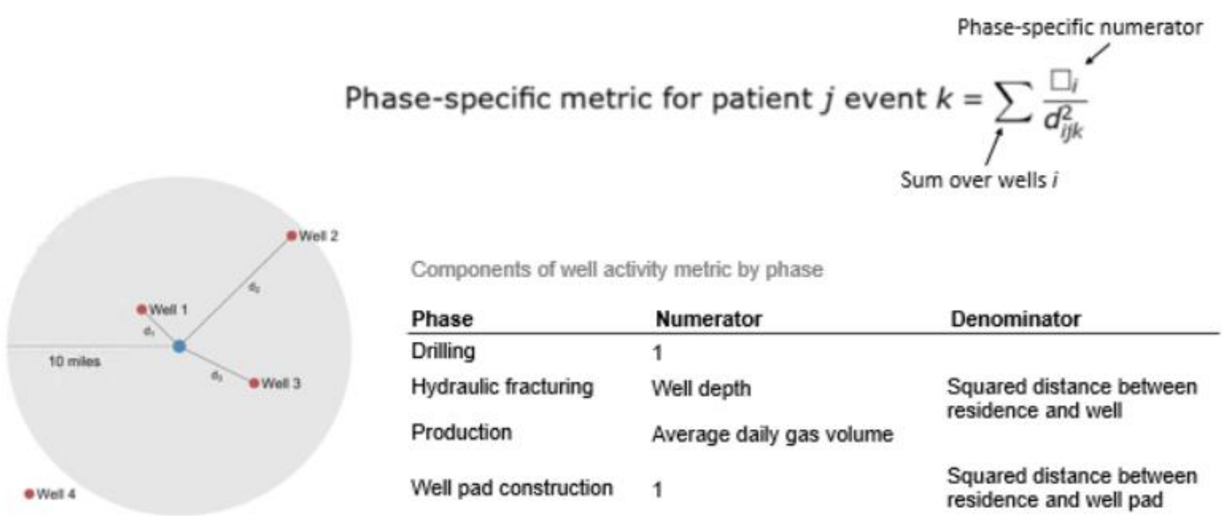

Figure 1. Well Phase Metric Calculation

We defined tertiles for each exposure metric (well pad construction phase, drilling phase, hydraulic fracturing phase, production phase) within each buffer distance (1, 2, 5, 10 miles):

- **Unexposed**: metric = 0
- **Exposed, low**: metric >0 and metric <33.3% of non-zero values among the entire cohort
- **Exposed, moderate**: metric >0 and metric ≥ 33.3% of non-zero values and metric <66.7% of non-zero values among the entire cohort
- **Exposed, high**: metric >0 and metric ≥ 66.7% of non-zero values among the entire cohort

Data Analysis

Data cleaning

We used graphical analyses, descriptive statistics, and exploratory data analysis to identify outlying observations, implausible values, and other inconsistencies, which were handled on a case-by-case basis, which occurred very infrequently. We examined all data for missingness. We computed the proportion of missing data for each variable contributing to the calculation of the exposure metric, the outcome variables, and the covariates. We stratified these calculations by year to examine patterns of missingness over time. We had no missing outcome information. If the proportions of missing covariate data were low (< 5%), we analyzed complete cases. We had greater than 5% unknown for BMI and smoking data. For BMI, we averaged BMI from the dates one year prior to the event date. Similarly, for smoking status, if a patient did not have a known smoking status on the event date, the most recent known smoking status prior to the event date was used.

For the UNGD exposure metric, we imputed missing well data using other available data. Missing well depths were imputed using the median well depth among wells not missing this measurement. Missing spud dates and stimulation dates were extrapolated using other available dates for each well and median phase durations among wells without missing dates.

Statistical analysis

We examined the four phases of the UNGD activity metric for correlation. In the event of substantial correlation among these four metrics, we would z-score each phase-specific metric, and then sum these z-scored phase-specific metrics to obtain a single, overall UNGD activity metric for each asthma exacerbation. However, we did not find evidence of correlation between the phases, and thus the phase-specific metrics were divided into three tertiles of exposure, representing low, moderate, and high UNGD activity, respectively.

We computed descriptive statistics (for continuous variables: mean and standard deviation or median and IQR; for categorical variables: frequency) for outcome variables and covariates. Descriptive statistics were calculated for each type of asthma exacerbation for cases and controls. We assessed differences in these distributions by running univariate logistic models using community as a random effect.

Our analyses assessed the association the phase-specific UNGD activity metrics (tertiles) with each of the three levels of asthma exacerbation severity. To do this, we fit a series of multilevel logistic regression models with a random intercept for community, as defined for the socioeconomic deprivation index, to account for nesting of patients within communities.

Each base model included all four phase-specific UNGD activity metrics. We then added to the base models: patient sex, year of encounter, race, family history of asthma, age category, smoking status, BMI category, season of event, type II diabetes diagnosis, community socioeconomic deprivation, and temperature (°C). We evaluated covariates for conditional significance (global tests assessing the covariate as a whole) using Wald or likelihood ratio tests . We also assessed trend for the tertiles of exposure using a Wald test for the linear form of the tertiles of exposure variables. We assessed multicollinearity among model covariates by calculating variance inflation factors (VIF).

Associations were reported as odds ratios comparing the tertile splits of the UNGD activity metric(s) to the unexposed group (reference level) with 95% confidence intervals. The odds ratio is used to determine whether a particular exposure (e.g., UNGD activity) is a risk factor for a particular type of asthma exacerbation, and to compare the magnitude of various risk factors for that outcome. Odds ratios (OR) can be interpreted as:

OR=1 Exposure (e.g., UNGD activity) does not affect odds of the type of asthma exacerbation

OR>1 Exposure (e.g., UNGD activity) is associated with higher odds of having the type of asthma exacerbation

OR<1 Exposure (e.g., UNGD activity) is associated with lower odds of having the type of asthma exacerbation

We used a two-sided type I error rate of 0.05 for significance testing. No adjustments were made for multiple comparisons. Analyses were performed using R version 4.1.2 (2021-11-01) and Stata 17 (StataCorp. 2021. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC). Forest plots were produced using GraphPad Prism version 9.5.1 for Windows (GraphPad Software, San Diego, California USA, www.graphpad.com). Forest plots are a graphical representation of odds ratios to facilitate comparisons across groups.

Results

Cohort Formation

Figure 3 shows the enrollment flowchart for the asthma cohort. We received 119,648 patients from R3, and our final cohort consisted of 46,676 patients.

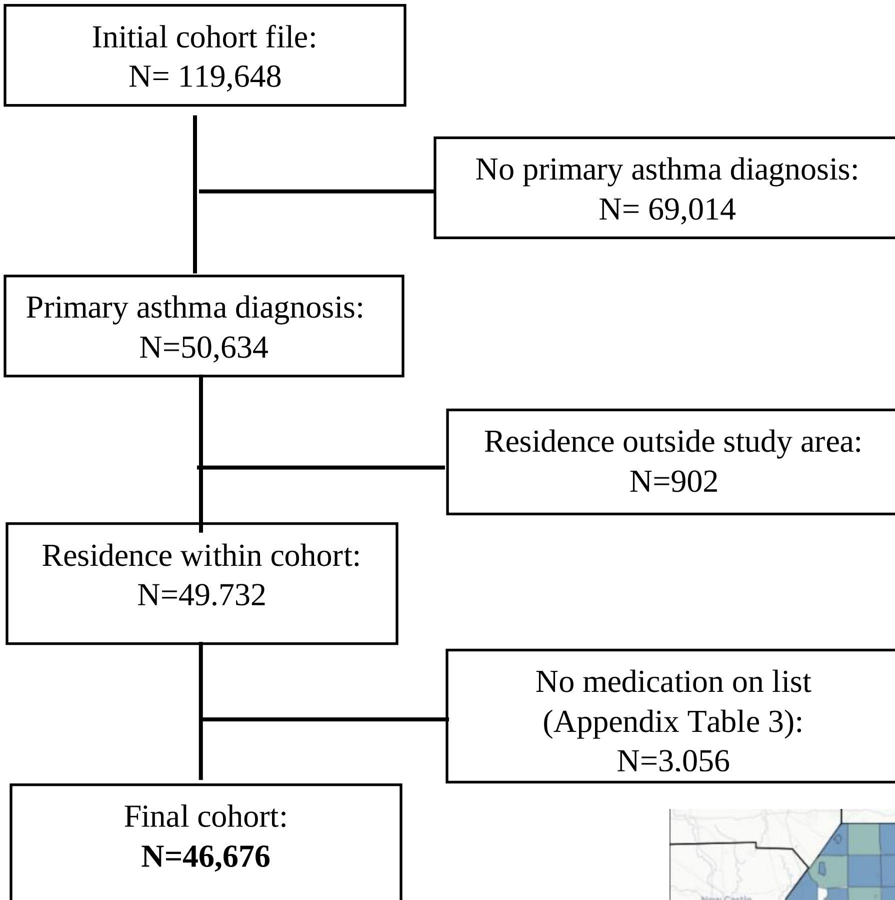

Figure 2. Cohort Enumeration Flowchart

The map (Figure 2) shows the counts for each patient community in the cohort. Allegheny County, excluding the City

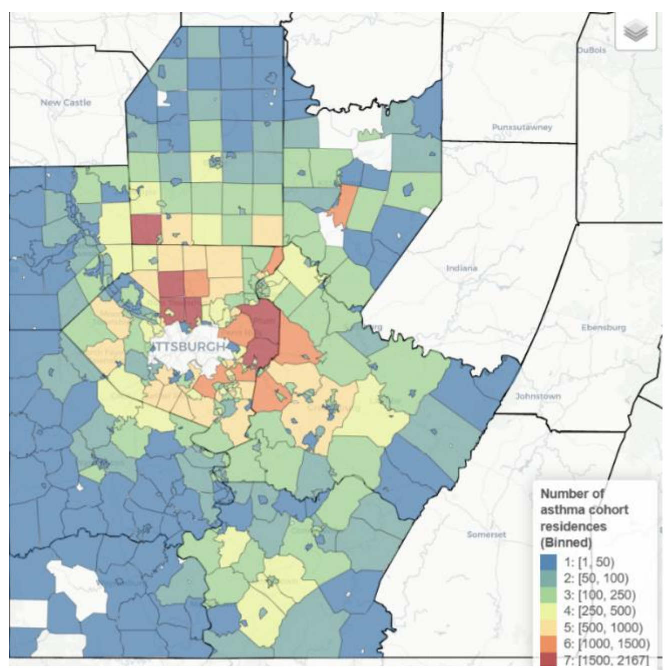

Figure 3. Map of Patient Communities

Events

One event per type, per year, per patient was randomly selected for our final set of case events. Table 3 shows the number of events per type and the number of controls. There were a total of 40,627 case and control events included.

Table 3. Final Counts by Event Type

Event Type	Number of Cases	Percent of Events	Number of Controls[1]
Severe	16,373	86.8	16,373
ED Severe	2,292	12.1	4,584
Hospitalization Severe	201	1.1	804

1- Control events frequency matched to case events by type: 1:1 severe, 2:1 ED severe; 4:1 hospitalization severe

Figure 4 shows the distribution of community socioeconomic deprivation index by quartile. Communities shown in blue are Quartile (Q) 1 (least deprivation) while communities shown in orange are in Q4 (most deprivation). Much of Allegheny County (excluding the City of Pittsburgh) and southern Butler County are in Q1.

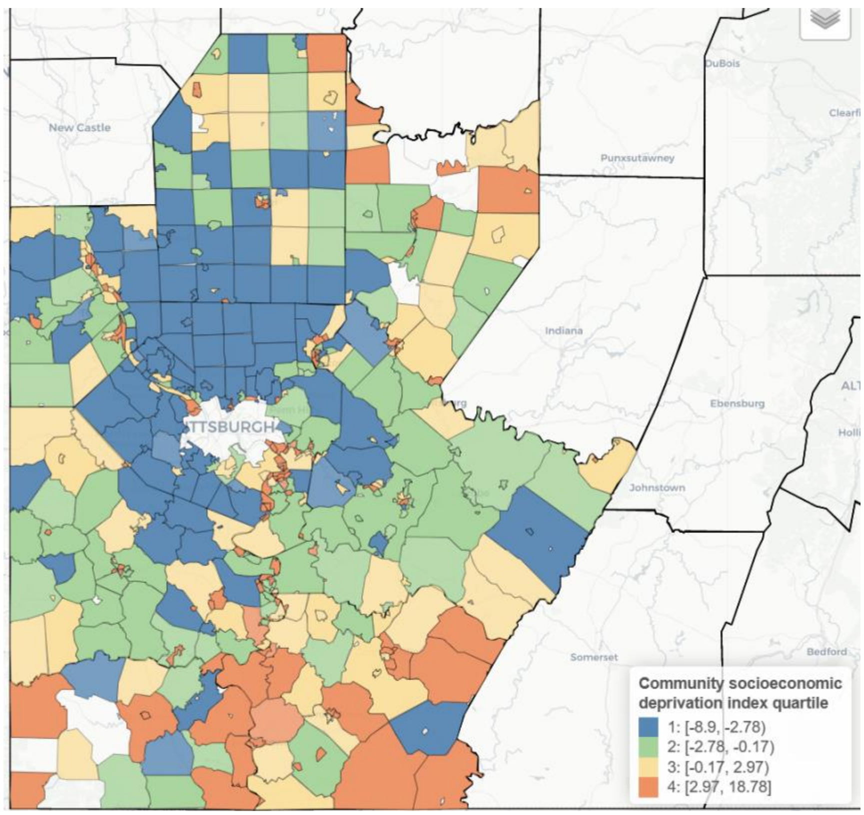

Figure 4. Map of Socioeconomic Deprivation Index by Community

UNGD Exposure

There were 5,799 wells included in our study from 2000 to 2020 (Figure 5). Through 2020, Washington County had the highest number of wells (n=1974), and Beaver County had the lowest number (n=141).

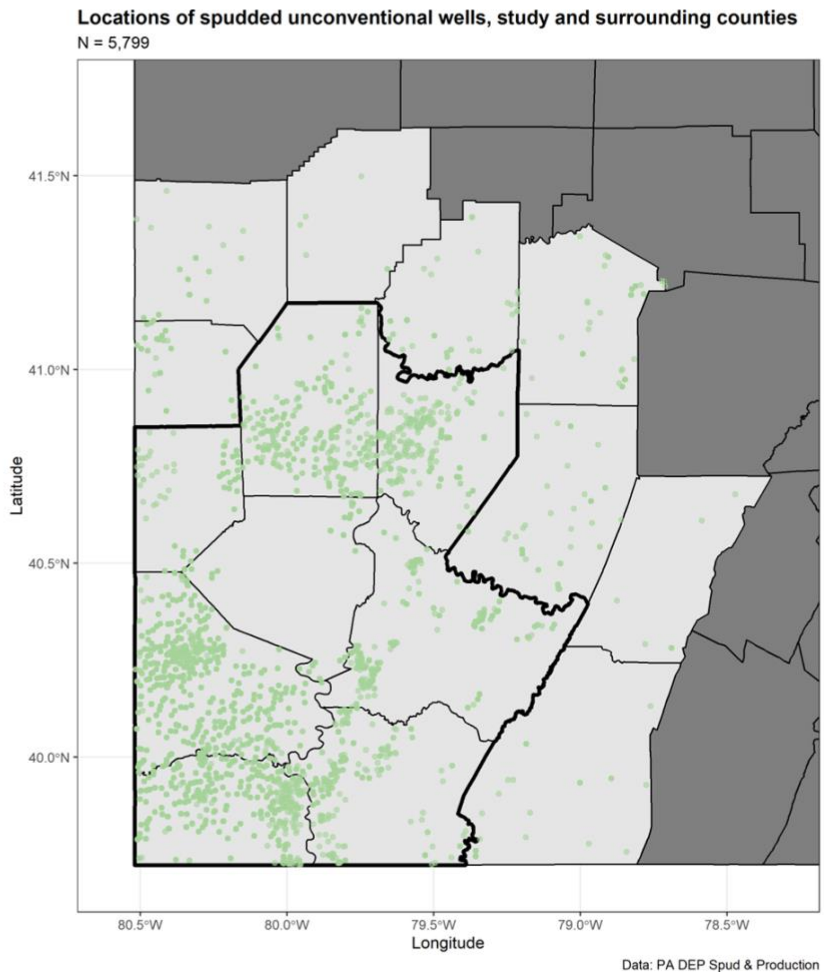

Figure 5. Map of UNGD Well Locations

There were fewer than 20 wells spudded in Southwestern Pennsylvania until 2007-2008, when production began increasing rapidly. The number of wells spudded peaked in 2014, with 765 as shown in Figure 6.

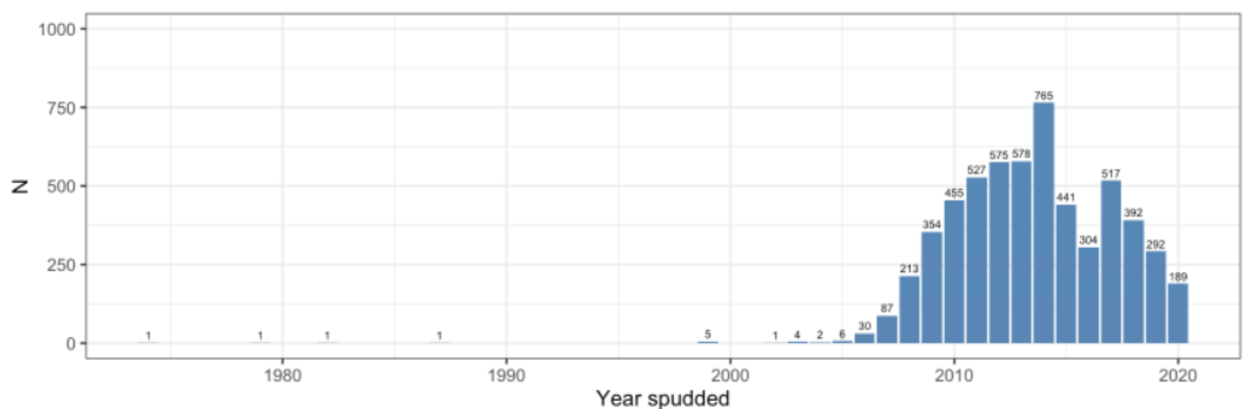

Figure 6. Histogram of UNGD Well Spud Dates by Year

Table 4 shows the median phase duration for each of the four UNGD activity metrics.

Table 4. UNGD Activity Metric Phase Durations

Phase	Phase name	Phase length
1	Well pad preparation	Minimum (spud date among wells on the pad) + 30 days 30 days
2	Drilling	Number of days between the spud and drilling completion dates Median: 104 days
3	Hydraulic fracturing	Number of days between stimulation commencement and stimulation completion Median: 12 days
4	Production	Duration of reporting period during which well reported production Mean: 2239 days (range 30-8769 days) Median: 2193 days

Shown in Table 5 are the cut points used for the tertiles (33.3% and 66.7%) as well as the minimum, median, and maximum value by phase and buffer. The production phase, which lasted the longest, had the highest metric values.

Table 5. Phase- and Buffer-Specific Cutpoints

Phase	Buffer (mi)	Min	33.3%	Median	66.7%	Max
Well pad preparation	1.0	3.88e-07	5.45e-07	7.96e-07	1.79e-06	7.31e-05
	2.0	9.70e-08	1.37e-07	1.75e-07	2.82e-07	7.31e-05
	5.0	1.54e-08	2.29e-08	3.02e-08	4.77e-08	7.32e-05
	10.0	3.86e-09	5.93e-09	8.38e-09	1.32e-08	7.32e-05
Drilling	1.0	3.94e-07	1.33e-06	2.05e-06	2.75e-06	3.09e-03
	2.0	9.65e-08	3.13e-07	5.02e-07	7.86e-07	3.09e-03
	5.0	1.54e-08	6.75e-08	1.08e-07	1.66e-07	3.09e-03
	10.0	3.86e-09	2.16e-08	3.56e-08	6.34e-08	3.09e-03
Hydraulic fracturing	1.0	1.38e-03	2.05e-03	2.32e-03	3.23e-03	2.15e-02
	2.0	2.10e-04	5.22e-04	7.08e-04	1.09e-03	2.15e-02
	5.0	3.05e-05	1.10e-04	1.72e-04	2.69e-04	2.15e-02
	10.0	6.47e-06	2.76e-05	4.23e-05	7.01e-05	2.15e-02
Production	1.0	1.65e-05	1.32e-02	4.46e-02	1.16e-01	3.73e+02
	2.0	3.69e-07	1.61e-02	3.30e-02	6.96e-02	3.73e+02
	5.0	7.23e-09	9.19e-03	2.40e-02	4.74e-02	3.73e+02
	10.0	6.62e-08	1.72e-02	1.72e-02	3.50e-02	3.73e+02

Characteristics by event type and case or control status are shown in Table 6, along with p-values assessing differences in distributions between the groups except for by sex, age or encounter year, which were matching variables. Severe exacerbations had statistically significant differences in distributions among cases and controls for all covariates. ED exacerbations had statistically significant differences in distributions among cases and controls for all covariates except family history of asthma and BMI. Hospitalizations had statistically significant differences in distributions among cases and controls in family history of asthma, BMI, and socioeconomic deprivation index.

Hospitalizations had the highest percentage of females. Severe exacerbations occurred most frequently among 5–13-year-olds, and ED and hospitalization exacerbations among 19–45-year-olds. Case events occurred most frequently in the winter and the majority of all patients were nonsmokers. More cases than controls were in Q1 (least deprived) and more controls than cases were in Q4 of the socioeconomic deprivation index for all event types.

Counts are also shown by event type by buffer distance, including for the 0.5-mile buffer. As shown, counts of exposed within the 0.5-mile buffer were so small as to preclude modeling. For all three event types, more case than control events were exposed for every well activity metric at every buffer distance for severe exacerbations. The greatest number of exposures occurred for the production phase among all events with up to 15% more case events exposed than control events.

Table 6. Descriptive Statistics of Cases and Controls by Asthma Exacerbation Type

	Severe Exacerbation		ED Exacerbation		Hospitalization Exacerbation	
Characteristic	Case, n=16,373 (%)	Control, n=16,373 (%)	Case, n=2292 (%)	Control, n=4584 (%)	Case, n=201 (%)	Control, n=804 (%)
Patient and Event Characteristics						
Female	9476 (57.9)	9476 (57.9)	1435 (62.6)	2870 (62.6)	141 (70.2)	564 (70.2)
Age in years, time of the event or matched encounter						
5 - <13	5065 (30.9)	5065 (30.9)	258 (11.3)	516 (11.2)	40 (19.9)	160 (19.9)
13 - <19	1710 (10.4)	1710 (10.4)	178 (7.8)	356 (7.8)	4 (2.0)	16 (2.0)
19 - <45	3425 (20.9)	3425 (20.9)	1048 (45.7)	2096 (45.7)	66 (32.8)	264 (32.8)
45 - <62	3533 (21.6)	3533 (21.6)	605 (26.4)	1210 (26.4)	52 (25.9)	208 (25.9)
62 - <75	2008 (12.3)	2008 (12.3)	172 (7.5)	344 (7.5)	29 (14.4)	116 (14.4)
75 - 90	632 (3.9)	632 (3.9)	31 (1.4)	62 (1.4)	10 (5.0)	40 (5.0)
Event year						
2011	1470 (9.0)	1470 (9.0)	93 (4.1)	186 (4.1)	11 (5.5)	44 (5.5)
2012	1625 (9.9)	1625 (9.9)	187 (8.2)	374 (8.2)	20 (9.9)	80 (9.9)
2013	1787 (10.9)	1787 (10.9)	216 (9.4)	432 (9.4)	20 (9.9)	80 (9.9)
2014	2168 (13.2)	2168 (13.2)	226 (9.9)	452 (9.9)	27 (13.4)	108 (13.4)
2015	2137 (13.1)	2137 (13.1)	234 (10.2)	468 (10.2)	26 (12.9)	104 (12.9)
2016	1640 (10.0)	1640 (10.0)	291 (12.7)	582 (12.7)	20 (9.9)	80 (9.9)
2017	1638 (10.0)	1638 (10.0)	257 (11.2)	514 (11.2)	22 (11.0)	88 (11.0)
2018	1502 (9.2)	1502 (9.2)	296 (12.9)	592 (12.9)	22 (11.0)	88 (11.0)
2019	1507 (9.2)	1507 (9.2)	324 (14.1)	648 (14.1)	18 (9.0)	72 (9.0)
2020	899 (5.5)	899 (5.5)	168 (7.3)	336 (7.3)	15 (7.5)	60 (7.5)
Family history of asthma (yes)	3044 (18.6)	2652 (16.2)	305 (13.3)	620 (13.5)	40 (19.9)	98 (12.2)
	<0.0001		0.772		0.007	

	Severe Exacerbation		ED Exacerbation		Hospitalization Exacerbation	
Characteristic	Case, n=16,373 (%)	Control, n=16,373 (%)	Case, n=2292 (%)	Control, n=4584 (%)	Case, n=201 (%)	Control, n=804 (%)
Race						
White	14,669 (89.6)	14,021 (85.6)	1881 (82.1)	3836 (83.7)	173 (86.1)	682 (84.8)
Black	1255 (7.7)	2, 011 (12.3)	202 (8.8)	661 (14.4)	24 (11.9)	105 (13.1)
Other/Unknown	448 (2.7)	351 (2.1)	209 (9.1)	87 (1.9)	4 (2.0)	17 (2.1)
	<0.0001		<0.0001		0.973	
Event season						
Winter: December 22 – March 21	4820 (29.4)	3884 (23.7)	688 (30.0)	1120 (24.4)	59 (29.4)	191 (23.8)
Spring: March 22 – June 21	3979 (24.3)	4045 (24.7)	533 (23.3)	1129 (24.6)	48 (23.9)	200 (24.9)
Summer: June 22 – September 21	2752 (16.8)	3990 (24.4)	402 (17.5)	1144 (25.0)	40 (19.9)	178 (22.1)
Fall: September 22 – December 21	4822 (29.5)	4454 (27.2)	669 (29.2)	1191 (26.0)	54 (26.9)	235 (29.2)
	<0.0001		<0.0001		0.416	
BMI						
Underweight or normal weight	5427 (33.1)	5744 (35.1)	636 (27.8)	1252 (27.3)	49 (24.4)	233 (29.0)
Overweight	3677 (22.5)	3559 (21.7)	613 (26.7)	1064 (23.2)	35 (17.4)	200 (24.9)
Obese	6852 (41.9)	6502 (39.7)	997 (43.5)	2141 (46.7)	111 (55.2)	351 (43.7)
Unknown	417 (2.5)	568 (3.5)	46 (2.0)	127 (2.8)	6 (3.0)	20 (2.5)
	<0.0001		0.126		0.014	
Smoking status						
Never	10,798 (65.9)	9922 (60.6)	1542 (67.3)	2518 (54.9)	122 (60.7)	450 (56.0)
Current	1466 (9.0)	1702 (10.4)	319 (13.9)	811 (17.7)	24 (11.9)	104 (12.9)
Former	2730 (16.7)	2827 (17.3)	396 (17.3)	957 (20.9)	43 (21.4)	180 (22.4)
Unknown	1379 (8.4)	1923 (11.7)	35 (1.5)	298 (6.5)	12 (6.0)	70 (8.7)
	<0.0001		<0.0001		0.480	
Community socioeconomic deprivation index, quartiles						
Q1	8875 (54.2)	8113 (49.6)	1315 (57.4)	2091 (45.6)	108 (53.7)	373 (46.4)
Q2	3365 (20.6)	3054 (18.7)	413 (18.0)	838 (18.3)	43 (21.4)	136 (16.9)
Q3	2083 (12.7)	2104 (12.8)	337 (14.7)	629 (13.7)	26 (12.9)	118 (14.7)
Q4	2050 (12.5)	3102 (18.9)	227 (9.9)	1026 (22.4)	24 (11.9)	177 (22.0)
	<0.0001		<0.0001		0.011	
Type II diabetes l (yes)	920 (5.6)	1599 (9.8)	78 (3.4)	515 (11.2)	25 (12.4)	78 (9.7)
	<0.0001		<0.0001		0.208	
Avg temperature day prior (°C) (SD)	14.8 (10.3)	16.8 (10.5)	15.1 (10.6)	16.8 (10.7)	15.7 (10.2)	16.2 (11.0)
	<0.0001		<0.0001		0.522	
Well Activity Metrics						
Exposed within 0.5-mile buffer						
Construction	8 (0.05)	7 (0.04)	1 (0.04)	1 (0.02)	0 (0.0)	0 (0.0)
Drilling	11 0.07)	3 (0.02)	4 (0.2)	1 (0.02)	0 (0.0)	0 (0.0)
Hydraulic fracturing	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Production	351 (2.1)	219 (1.3)	37 (1.6)	55 (1.2)	2 (1.0)	11 (0.01)
Exposed within 1 mile buffer						
Construction	24 (0.1)	18 (0.1)	4 (0.2)	1 (0.02)	1 (0.5)	0 (0.0)
Drilling	77 (0.5)	41 (0.3)	10 (0.4)	9 (0.2)	1 (0.5)	0 (0.0)
Hydraulic fracturing	9 (0.05)	5 (0.03)	1 (0.04)	1 (0.02)	0 (0.0)	0 (0.0)
Production	1131 (6.9)	842 (5.1)	119 (5.2)	211 (4.6)	11 (5.5)	30 (3.7)
Exposed within 2-mile buffer						
Construction	98 (0.6)	70 (0.4)	6 (0.3)	11 (0.2)	1 (0.5)	2 (0.2)
Drilling	369 (2.2)	262 (1.6)	42 (1.8)	55 (1.2)	3 (1.5)	11 (1.4)
Hydraulic fracturing	70 (0.4)	62 (0.4)	6 (0.3)	11 (0.2)	2 (1.0)	1 (0.1)

	Severe Exacerbation		ED Exacerbation		Hospitalization Exacerbation	
Characteristic	Case, n=16,373 (%)	Control, n=16,373 (%)	Case, n=2292 (%)	Control, n=4584 (%)	Case, n=201 (%)	Control, n=804 (%)
Production	3122 (19.1)	2270 (13.9)	368 (16.1)	637 (13.9)	32 (15.9)	107 (13.3)
Exposed within 5-mile buffer						
Construction	683 (4.2)	516 (3.1)	73 (3.2)	117 (2.6)	8 (4.0)	18 (2.2)
Drilling	2496 (15.2)	1883 (11.5)	302 (13.2)	512 (11.2)	28 (14.0)	74 (9.2)
Hydraulic fracturing	652 (4.0)	491 (3.0)	78 (3.4)	111 (2.4)	5 (2.5)	19 (2.4)
Production	8646 (52.8)	6853 (41.9)	1200 (52.3)	1890 (41.2)	99 (49.2)	317 (39.4)
Exposed within 10-mile buffer						
Construction	2706 (16.5)	2305 (14.1)	336 (14.7)	537 (11.7)	34 (16.9)	101 (12.6)
Drilling	7974 (48.7)	6937 (42.4)	1032 (45.0)	1841 (40.1)	89 (44.3)	312 (38.8)
Hydraulic fracturing	2937 (17.9)	2391 (14.6)	386 (16.8)	610 (13.3)	29 (14.4)	110 (13.7)
Production	14,825 (90.5)	13,133 (80.2)	2075 (90.5)	3686 (80.4)	183 (91.0)	625 (77.7)

Model Results

Severe Exacerbation Models

Adjusted models for severe asthma exacerbations are shown below. For the construction, drilling, and hydraulic fracturing phases, there were no consistent associations at any buffer distance. For the production phase, there were statistically significantly elevated odds ratios of 3 to 5 for all buffer distances, some of which increased with increasing intensity of exposure. For all buffer distances, both the global and trend p-values were statistically significant.

Table 7. Asthma Severe Exacerbation Models by Well Phase Activity Metric

Buffer	Adjusted OR[1] (95% CI)			
	Well Pad Preparation	Drilling	Hydraulic Fracturing	Production
1 mile				
Unexposed	--	--	--	--
Low	1.50 [0.53, 4.25]	1.81 [0.92, 3.54]	0.94 [0.15, 5.88]	3.80 [3.09, 4.67]**
Moderate	0.96 [0.29, 3.13]	1.36 [0.70, 2.65]	3.44 [0.37, 32.20]	3.83 [3.13, 4.67]**
High	0.65 [0.21, 1.95]	1.58 [0.77, 3.23]	0.87 [0.12, 6.30]	3.81 [3.11, 4.66]**
Global p-value	0.87	0.02	0.04	<0.0001
Trend p-value	0.07	<0.0001	<0.0001	<0.0001
2 miles				
Unexposed	--	--	--	--
Low	1.49 [0.84, 2.65]	1.01 [0.75, 1.36]	0.62 [0.35, 1.11]	4.52 [3.89, 5.25]**
Moderate	0.55 [0.32, 0.94]	1.22 [0.91, 1.63]	1.11 [0.59, 2.10]	5.12 [4.41, 5.95]**
High	1.11 [0.63, 1.96]	1.05 [0.79, 1.41]	0.98 [0.51, 1.90]	4.02 [3.45, 4.67]**
Global p-value	0.15	0.14	0.02	<0.0001
Trend p-value	0.09	<0.0001	0.001	<0.0001
5 miles				
Unexposed	--	--	--	--
Low	1.03 [0.83, 1.27]	1.12 [0.99, 1.26]	1.17 [0.94, 1.45]	4.41 [3.92, 4.96]**
Moderate	1.00 [0.82, 1.23]	1.10 [0.98, 1.24]	1.06 [0.86, 1.32]	4.63 [4.10, 5.24]**
High	0.93 [0.75, 1.14]	1.05 [0.93, 1.19]	0.99 [0.80, 1.22]	4.73 [4.14, 5.39]**
Global p-value	0.96	0.25	0.14	<0.0001
Trend p-value	0.36	0.01	0.22	<0.0001
10 miles				
Unexposed	--	--	--	--
Low	0.97 [0.87, 1.08]	1.02 [0.95, 1.10]	1.04 [0.93, 1.15]	3.53 [3.20, 3.89]**
Moderate	1.05 [0.95, 1.17]	1.07 [1.00, 1.15]	1.11 [1.00, 1.23]	4.29 [3.85, 4.78]**
High	0.99 [0.88, 1.10]	1.07 [0.99, 1.16]	1.10 [0.99, 1.22]	4.72 [4.18, 5.34]**
Global p-value	0.69	0.21	0.11	<0.0001
Trend p-value	0.52	0.12	0.10	<0.0001

1- Models adjusted for SES, encounter year, age category, sex, race, season, BMI category, smoking status, family history of asthma, temperature, and history of type II diabetes

* p<0.05; ** p<0.001

The severe exacerbation forest plots by buffer distance for each phase are shown in Figure 7. The vertical line at 1 represents a null relationship; dots below 1 indicate reduced risk and dots above 1 indicate increased risk.

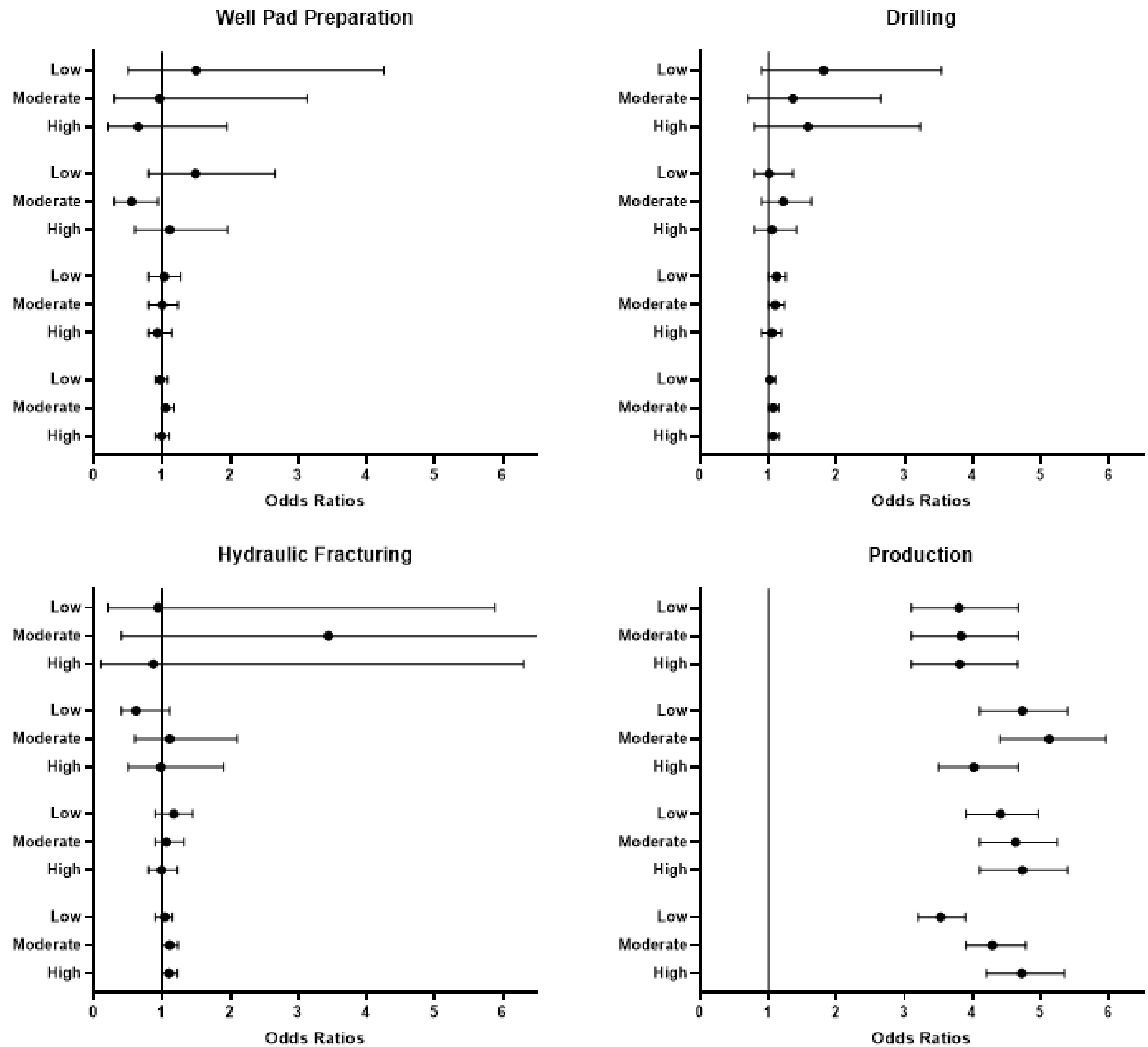

Figure 7. Forest Plots of Model Results for Severe Exacerbations

ED Severe Exacerbation Models

The adjusted models for exacerbations requiring an ED visit are shown in Table 8. Some of the exposure characterizations, noted as Not Applicable (NA), could not be modeled due to the number of cases and controls within the smaller buffer distances for this outcome (see Table 5 for counts of exposed cases and controls at each buffer distance). For the construction, drilling, and hydraulic fracturing phases, there were no consistent associations at any buffer distance. For the production phase, there were statistically significantly elevated odds ratios between 2 and 6 for all buffer distances, most of which increased with increasing intensity of exposure. For all buffer distances, both the global and trend p-values were statistically significant.

Table 8. Asthma ED Severe Exacerbation Model Results by Well Phase Activity Metric

Buffer	Adjusted OR[1] (95% CI)			
	Well Pad Preparation	Drilling	Hydraulic Fracturing	Production
1 mile				
Unexposed				--
Low				3.44 [1.85, 6.40]**
Moderate	NA[2]	NA	NA	3.96 [2.28, 6.87]**
High				4.86 [2.90, 8.16]**
Global p-value				<0.0001
Trend p-value				<0.0001
2 miles				
Unexposed		--		--
Low		0.98 [0.43, 2.24]		3.42 [2.33, 5.03]**
Moderate	NA	1.63 [0.69, 3.85]	NA	3.41 [2.37, 4.92]**
High		1.19 [0.49, 2.87]		4.13 [2.82, 6.05]**
Global p-value		0.74		<0.0001
Trend p-value		0.22		<0.0001
5 miles				
Unexposed	--	--	--	--
Low	0.87 [0.51, 1.50]	0.81 [0.60, 1.09]	0.98 [0.54, 1.79]	4.89 [3.65, 6.54]**
Moderate	1.55 [0.86, 2.81]	0.93 [0.67, 1.27]	1.17 [0.68, 2.01]	5.01 [3.71, 6.78]**
High	0.72 [0.36, 1.43]	0.90 [0.63, 1.27]	0.84 [0.41, 1.71]	4.11 [2.96, 5.70]**
Global p-value	0.31	0.71	0.77	<0.0001
Trend p-value	0.16	0.49	0.62	<0.0001
10 miles				

Buffer	Adjusted OR[1] (95% CI)			
	Well Pad Preparation	Drilling	Hydraulic Fracturing	Production
Unexposed	--	--	--	--
Low	0.96 [0.72, 1.28]	1.08 [0.91, 1.30]	0.93 [0.70, 1.24]	3.50 [2.75, 4.45]**
Moderate	1.32 [0.98, 1.77]	0.87 [0.72, 1.05]	1.23 [0.95, 1.61]	4.49 [3.45, 5.84]**
High	1.13 [0.83, 1.53]	0.82 [0.67, 1.01]	1.09 [0.82, 1.45]	4.81 [3.58, 6.47]**
Global p-value	0.28	0.07	0.39	<0.0001
Trend p-value	0.10	0.07	0.35	<0.0001

1- Models adjusted for SES, exposure year, age category, sex, race, season, BMI category, smoking status, family history of asthma, temperature, and history of type II diabetes

2- Small sample sizes precluded modeling

* p<0.05; ** p<0.001

Figure 8 shows the ED exacerbation forest plots by buffer distance for each phase. The vertical line at 1 represents a null relationship; dots below 1 indicate reduced risk and dots above 1 indicate increased risk.

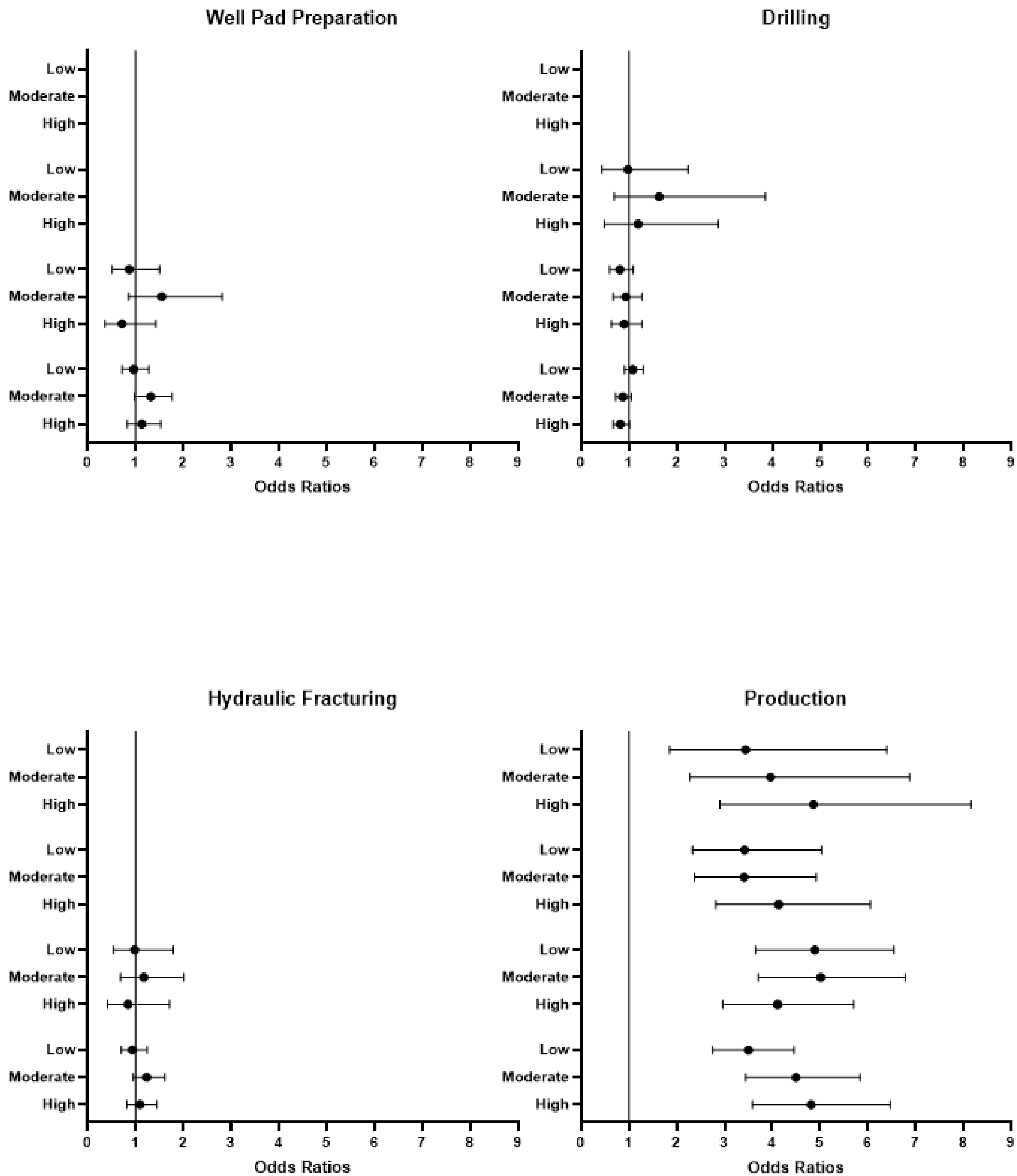

Figure 8. Forest Plots of Model Results for Emergency Department Severe

Hospitalization Severe Exacerbation Models

Adjusted model results for events requiring hospitalization are shown in Table 9. Some of the exposure characterizations, noted as Not Applicable (NA), could not be modeled due to the smaller number of cases (n=201) and controls (n=804) for this outcome (see Table 5 for counts of exposed cases and controls at each buffer distance). Only production could be modeled at the 1- and 2-mile buffers. For the construction, drilling, and hydraulic fracturing phases, there were no consistent associations at any buffer distance. For the production phase, all odds ratios were elevated and those odds ratios from 3 to 8 were statistically significantly. Most of the odds ratios increased with increasing intensity of exposure. For all buffer distances, both the global and trend p-values were statistically significant.

Table 9. Asthma Hospitalization Severe Exacerbation Model Results by Well Phase Activity Metric

Buffer	Adjusted OR[1] (95% CI)			
	Well Pad Preparation	Drilling	Hydraulic Fracturing	Production
1 mile				
Unexposed				--
Low				1.58 [0.30, 8.32]
Moderate	NA[2]	NA	NA	4.08 [1.01, 16.48]*
High				6.89 [1.54, 30.89]*
Global p-value				0.001
Trend p-value				0.001
2 miles				
Unexposed				--
Low				2.01 [0.77, 5.26]
Moderate	NA	NA	NA	2.33 [0.83, 6.55]
High				8.71 [3.09, 24.55]**
Global p-value				0.0001
Trend p-value				0.01
5 miles				
Unexposed	--	--		--
Low	1.59 [0.35, 7.15]	0.67 [0.24, 1.86]		3.68 [1.79, 7.59]**
Moderate	0.93 [0.15, 5.79]	1.32 [0.52, 3.34]	NA	3.08 [1.48, 6.42]*
High	1.75 [0.29, 10.57]	1.57 [0.66, 3.76]		4.77 [2.18, 10.45]**
Global p-value	0.95	0.58		0.0007
Trend p-value	0.52	0.40		0.01
10 miles				
Unexposed	--	--	--	--

Buffer	Adjusted OR[1] (95% CI)			
	Well Pad Preparation	Drilling	Hydraulic Fracturing	Production
Low	1.25 [0.61, 2.54]	0.85 [0.50, 1.45]	0.94 [0.44, 2.01]	3.13 [1.69, 5.81]**
Moderate	0.95 [0.42, 2.15]	0.64 [0.37, 1.13]	0.83 [0.37, 1.82]	3.64 [1.87, 7.09]**
High	0.83 [0.36, 1.87]	1.44 [0.81, 2.55]	0.49 [0.19, 1.26]	4.64 [2.25, 9.58]**
Global p-value	0.88	0.10	0.51	0.0003
Trend p-value	0.94	0.98	0.18	<0.0001

1- Models adjusted for SES, exposure year, age category, sex, race, season, BMI category, smoking status, family history of asthma, temperature, and history of type II diabetes

2- Small sample sizes precluded modeling

* p<0.05; ** p<0.001

The hospitalization severe exacerbation forest plots by buffer distance for each phase are shown in Figure 9. The vertical line at 1 represents a null relationship; dots below 1 indicate reduced risk and dots above 1 indicate increased risk.

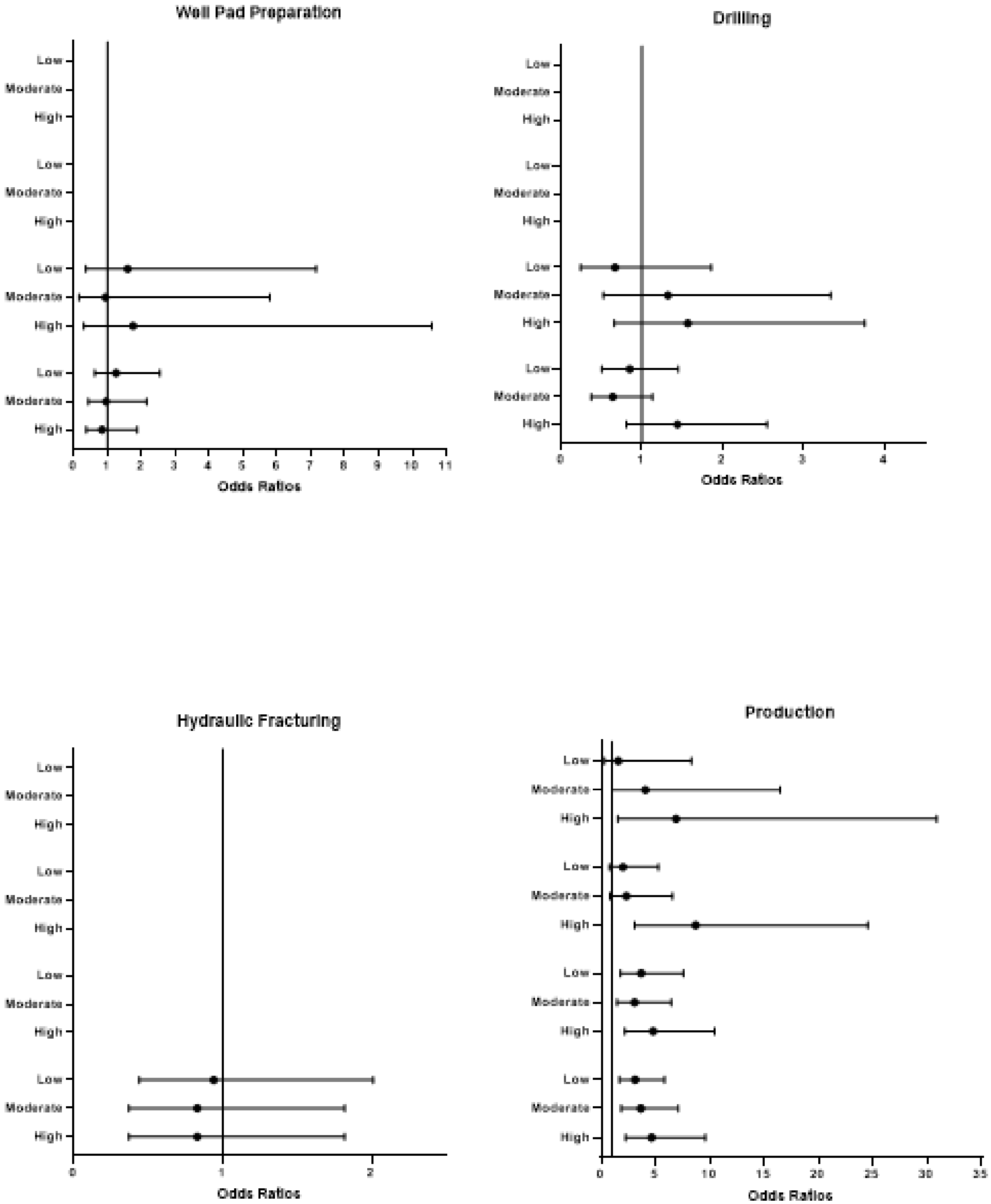

Figure 9. Forest Plots for Model Results for Hospitalization Severe Exacerbations

Discussion and Conclusions

This study examined three types of asthma events, severe, ED severe, and hospitalization severe, among more than 40,000 patients in an eight-county area of Southwestern PA from 2011-2020. To help frame the study conclusions, we are using the following classifying terms and criteria:

1. There are no data to suggest/support an increased risk:
   a. No statistically significantly elevated odds ratios
   b. Odds ratios at or near 1
   c. Odds ratios below 1 (with or without statistical significance)

2. There are limited data to suggest/support an increased risk:
   a. Statistically significantly elevated odds ratios in a low or moderate tertile
   b. Not statistically significant elevated odds ratios in multiple tertiles

3. There are moderate data to support an increased risk:
   a. Statistically significantly elevated odds ratios in multiple low or moderate tertiles
   b. Statistically significantly elevated odds ratios in a high tertile

4. There are strong data to suggest/support an increased risk:
   a. Statistically significantly elevated odds ratios in multiple tertiles
   b. Statistically significantly elevated odds ratios that increase across low, moderate, and high tertiles

We found strong evidence to suggest an increased risk in the production phase for all buffer distances examined for all three event types, based on consistent, statistically significantly elevated odds ratios. Elevations ranged from 2 to 8 times the baseline of no wells within 10 wells of the patient residence.

For all three event types, there were no data to support an increased risk at any buffer distance for the well pad preparation, drilling, and hydraulic fracturing phases.

This study replicated earlier work in Northeastern PA by Rasmussen et al[8]. In that study, they did not enforce buffer distances in their well activity metrics (all wells were included). Therefore, the most applicable comparison to these results is using our 10-mile buffer distances. Table 10 shows the odds ratios from this study compared to those from Rasmussen. The Rasmussen study found elevations in the well pad preparation, drilling, and hydraulic fracturing phases that were not found in this study.

Conversely, compared to Rasmussen, this study found much higher odds ratios for the production phase. Rasmussen reported an odds ratio of 4.43 in their highest tertile of production for their equivalent of our severe exacerbation, an order for an oral corticosteroid (OCS). That is similar to the odds ratio of 4.72 reported in this study. However, this study found elevated odds ratios for all tertiles of all buffers in the production phase.

Of note is that this study found the highest odds ratios for all asthma endpoints during the production phase. This could suggest that this phase may represent unique exposures not encountered during other phases. For example, this phase might be associated with more natural gas- and shale-derived hydrocarbons as well as produced water and, perhaps to some extent, hydraulic fracking fluid flowback. Moreover, this phase is generally the longest phase of well development and, thus, provides greater opportunity for chronic and cumulative exposures.

While this study had many similarities to that of Rasmussen et al, there are some notable differences. Rasmussen used a slightly less conservative definition for severe exacerbations of a new OCS medication ordered. We used the current ATS recommended definition of an initiation or increase of OCS medication in our three types of events. This could have led to less severe exacerbations included in the Rasmussen study than in ours.

Rasmussen et al. included events from 2005 to 2012 compared to 2011 to 2020 in our study. The 2011-to-2020-time frame was a particularly active time for UNGD development in Southwestern PA, but also encompassed technological advancements which may have modified exposure over time. We included encounter year as a covariate in our models to help account for these changes. Additionally, Rasmussen et al. did not enforce a buffer but included all wells in PA in their activity metrics, while this study specifically investigated the impacts at various buffer distances and excluded wells further than 10 miles from metric calculations.

Each study used electronic health records from a large provider in their region. The demographic characteristics are similar in Northeastern PA and our eight-county region in Southwestern PA, particularly with the exclusion of the City of Pittsburgh. However, Rasmussen et al. had a higher proportion of white patients in all case and control groups for all event types except ED cases. This study had a higher proportion of patients 5 to 13 years old than did Rasmussen et al. The high proportion of events among younger patients provides additional support that these are asthma exacerbations and not due to a chronic condition affecting older patients, such as chronic pulmonary obstructive disorder.

Strikingly, while our overall sample sizes were similar (n=46,676 patients in this study; n=35,508 patients in Rasmussen et al.), the Rasmussen et al. study had a much higher proportion of hospitalizations (n=4782 case events compared to n=201 case events in this study). This could be an indication of more poorly controlled asthma in that population which led to a higher proportion of very severe events. While we do not have similar information from Rasmussen, all of our hospitalization cases also had severe or ED exacerbations; there were no patients who only had a hospitalization during this timeframe. Among our severe exacerbations, only 7% (n=1122) had an ED or hospital exacerbation and 31.5% of our ED cases had severe (primarily) or hospitalization (rarely) exacerbations. This provides additional support that these findings are robust and are not being driven by a small number of patients with multiple endpoints.

Table 10. Comparison of Adjusted Odds Ratios in Current Study[1] with those in Rasmussen et al. (2016)[2]

	Severe Exacerbations		ED Exacerbations		Hospitalizations	
	Pitt SPH 10-mile buffer	Rasmussen 2016 (OCS[3] Orders)	Pitt SPH 10-mile buffer	Rasmussen 2016	Pitt SPH 10-mile buffer	Rasmussen 2016
Well Pad Preparation						
Unexposed	--	--	--	--	--	--
Low	0.97 (0.87, 1.08)	1.54 (1.37-1.74)*	0.87 (0.51, 1.50)	1.53 (1.06-2.23)*	1.25 (0.61, 2.54)	1.26 (1.06-1.50)*
Moderate	1.05 (0.95, 1.17)	1.66 (1.47-1.87)*	1.55 (0.86, 2.81)	1.77 (1.20-2.60)*	0.95 (0.42, 2.15)	1.37 (1.15-1.64)*
High	0.99 (0.88, 1.10)	1.59 (1.41-1.81)*	0.72 (0.36, 1.43	1.37 (0.94-1.99)	0.83 (0.36, 1.87)	1.45 (1.21-1.73)*
Drilling						
Unexposed	--	--		--	--	--
Low	1.02 (0.95, 1.10)	1.45 (1.29-1.63)*	0.81 (0.60, 1.09)	1.53 (1.06-2.21)*	0.85 (0.50, 1.45)	1.16 (0.98-1.37)
Moderate	1.07 (1.00, 1.15)	1.45 (1.29-1.63)*	0.93 (0.67, 1.27)	1.54 (1.04-2.27)*	0.64 (0.37, 1.13)	1.26 (1.05-1.50)*
High	1.07 (0.99, 1.16)	1.99 (1.75-2.26)*	0.90 (0.63, 1.27)	1.57 (1.08-2.29)*	1.44 (0.81, 2.55)	1.64 (1.38-1.97)*
Hydraulic Fracturing						
Unexposed	--	--		--	--	--
Low	1.04 (0.94, 1.15)	1.23 (1.09-1.39)*	0.93 (0.70, 1.24)	1.51 (1.05-2.19)*	0.94 (0.44, 2.01)	1.13 (0.96-1.33)
Moderate	1.11 (1.00, 1.23)	2.22 (1.95-2.53)*	1.23 (0.95, 1.61)	1.74 (1.17-2.61)*	0.83 (0.37, 1.82)	1.31 (1.10-1.57)*
High	1.10 (0.99, 1.22)	3.00 (2.60-3.45)*	1.09 (0.82, 1.45)	1.71 (1.16-2.52)*	0.83 (0.37, 1.82)	1.66 (1.38-1.98)*
Production						
Unexposed	--	--	--	--	--	--
Low	3.53 (3.20, 3.89)*	1.28 (1.13-1.46)*	3.50 (2.75, 4.45)*	1.47 (1.01-2.14)*	3.13 (1.69, 5.81)*	1.10 (0.92-1.30)
Moderate	4.29 (3.85, 4.78)*	2.15 (1.87-2.47)*	4.49 (3.45, 5.84)*	1.10 (0.74-1.65)	3.64 (1.87, 7.09)*	1.16 (0.97-1.38)
High	4.72 (4.18, 5.34)*	4.43 (3.75-5.22)*	4.81 (3.58, 6.47)*	2.19 (1.47-3.25)*	4.64 (2.25, 9.58)*	1.74(1.45-2.09)*

1 - Models adjusted for SES, exposure year, age category, sex, race, season, BMI category, smoking status, family history of asthma, temperature, and history of type II diabetes

2- From Rasmussen: Multilevel models with a random intercept for patient and community were adjusted for age category (5-12, 13-18, 19-44, 45-61, 62-74, >=75 years), sex (male or female), race/ethnicity (white, black, Hispanic, or other), family history of asthma (yes vs no), smoking status (never, former, current, or missing), season (spring, March 22–June 21; summer, June 22–September 21; fall, September 22–December 21; winter, December 22–March 21), Medical Assistance (yes vs no), overweight/obesity status (normal, body mass index [BMI], <85th percentile for children or <25 for adults; overweight, BMI, 85th to <95th percentile for children or 25 to <30 for adults; obese, BMI, >=95th percentile for children or>=30 for adults; or BMI missing), type 2 diabetes (yes vs no), community socioeconomic deprivation (across quartiles), distance to nearest major and minor arterial road (truncated at the 98th percentile, measured in meters, z transformed), squared distance to nearest major and minor arterial road (truncated at the 98th percentile, measured in meters, z transformed), maximum temperature on the day prior to event (measured in degrees Celsius), and squared maximum temperature on the day prior to event (measured in degrees Celsius).

3- Oral corticosteroids

* Statistically significant

Strengths and Limitations

This study has many strengths, including case ascertainment from a large health system with a large footprint in Southwestern Pennsylvania. However, we may have missed patients who used other health systems or facilities outside of this network for their care. We had few patients from Greene County; although this is the least populated county within our study area, it could also indicate that residents are receiving care outside of this network, including in neighboring West Virginia. We relied on electronic health records for our cohort information, which may not be reliable for some of our covariates, including but not limited to race and smoking status. These records may also fail to completely capture family history of asthma, and using ICD codes may not fully capture all cases of diabetes. This identification could be improved by including blood sugar and medication information. Additionally, individuals who do not have private insurance and those with more limited access to care could indicate a referral bias. This may partially explain the statistically significant differences among cases and controls for the socioeconomic deprivation index; there were fewer cases than controls for each event type in Quartile 4 (most deprivation).

The study applied a rigorous well phase activity assessment using multiple buffers to assess the strength of associations - the first to do so. These phase-by-buffer analyses provide new and important information about the associations of UNGD and asthma exacerbations. However, even in our large, system-based cohort, we had small sample sizes in some analyses, especially those within our smallest buffer distances and during shorter well activity phases (e.g., hydraulic fracturing). Additionally, we did not adjust for multiple comparisons. Some of the relationships between outcome and exposure may indicate evidence of a threshold effect, which was not assessed in the functional forms of the exposures examined here. Future studies should examine non-linear and other functional forms. The trend test assessed the linear relationship of the exposure tertiles, and some trend tests were statistically significant even when odds ratios (or term birthweights) were close to the reference level. Our well phase activity metric does not directly assess exposures to specific hazards associated with UNGD activity. The drop in cases in 2020 may indicate that we did not have complete coverage in that year but could also be an impact of the Covid-19 pandemic. However, our more than 10-year examination of cases lends additional credibility to these results. Additionally, the geocoding restrictions may have impacted exposure assignments at small buffers; however, we do not anticipate that this non-differential misclassification would have influenced the results.

This study provides evidence of associations between UNGD and asthma exacerbations. Future analyses should consider a more direct exposure pathway than our UNGD metric. These results should also be examined by age group to understand whether those most vulnerable, including children and the elderly, are more strongly impacted. Additionally, we considered exposures only one day prior to the event. Other windows, including those from 2-5 days prior, should be examined to ensure the effects are similar.

Our UNGD exposure metric was based on residence in the electronic health records. However, exposures occur outside of the home as well, including at daycare, school, and work. Future work should consider the impact of these non-residential exposures as well. Additionally, as our buffer distances increased, the opportunity for non-well exposures increased. Asthma exacerbations could be associated with other additional exposures that may influence air quality, such as UNGD infrastructure and non-UNGD exposures.

Appendix

Table A1. List of zip codes located all or in part in the City of Pittsburgh in Allegheny County

Zip code	All or part City of Pittsburgh
15106	Part City
15120	Part City
15201	All City
15203	All City
15204	Part City
15205	Part City
15206	All City
15207	All City
15208	All City
15210	Part City
15211	All City
15212	Part City
15213	All City
15214	Part City
15215	Part City
15216	Part City
15217	All City
15218	Part City
15219	All City
15220	Part City
15221	Part City
15222	All City
15224	All City
15226	Part City
15227	Part City
15230	All City
15232	All City
15233	All City
15234	Part City
15235	Part City
15240	Part City
15260	All City
15282	All City

Table A2. **Study population inclusion and exclusion criteria ICD-9 and ICD-10 codes**

Name	ICD-9 codes	ICD-10 codes
Inclusion criteria		
Asthma	493.00, 493.01, 493.02, 493.10, 493.11, 493.12, 493.20, 493.21, 493.22, 493.81, 493.82, 493.90, 493.91, 493.92	J45.20, J45.22, J45.21, J45.990, J45.991, J45.909, J45.998, J45.902, J45.901
Exclusion criteria		
Cystic fibrosis	277.00, 277.01, 277.02, 277.03, 277.09	E84.9, E84.11, E84.0, E84.19, E84.8
Chronic pulmonary heart disease	416.0, 416.1, 416.2, 416.8, 416.9	I27.0, I27.1, 127.82, I27.2, I27.89, I27.81, I27.9
Paralysis of vocal cords or larynx	478.30, 478.31, 478.32, 478.33, 478.34	J38.00, J38.01, J38.02
Bronchiectasis	494.0, 494.1	J47.9, J47.1
Pneumoconiosis	500, 501, 502, 503, 504, 505, 506.0, 506.1, 506.2, 506.3, 506.4, 506.9, 507.0, 507.1, 507.8, 508.0, 508.1, 508.2, 508.8, 508.9	J60, J61, J62.8, J63.0, J63.1, J63.2, J63.3, J63.4, J63.5, J63.6, J66.0, J66.1, J66.2, J66.8, J64, J68.0, J68.1, J68.2, J68.3, J68.4, J68.9, J69.0, J69.1, J69.8, J70.0, J70.1, J70.5, J70.8, J70.9

Table A3. Oral corticosteroid medication order exclusion criteria ICD-9 and ICD-10 codes

Name	ICD-9 codes	ICD-10 codes
Suppurative and unspecified otitis media	382.00, 382.01, 382.02, 382.1, 382.2, 382.3, 382.4, 382.9	H66.009, H66.019, H67.9, H66.13, H66.23, H66.3X9, H66.40, H66.90
Non-suppurative otitis media and Eustachian tube disorders	381.00, 381.01, 381.02, 381.03, 381.04, 381.05, 381.06, 381.10, 381.19, 381.20, 381.29, 381.3, 381.4, 381.50, 381.51, 381.52, 381.60, 381.61, 381.62, 381.63, 381.7, 381.81, 381.89, 381.9	H65.199, H65.00, H65.119, H65.20, H65.30, H65.499, H65.90, H68.009, H68.019, H68.029, H68.109, H68.119, H68.129, H68.139, H69.00, H69.80, H69.90
Contact dermatitis and other eczema	692.0, 692.1, 692.2, 692.3, 692.4, 692.5, 692.6, 692.70, 692.71, 692.72, 692.73, 692.74, 692.75, 692.76, 692.77, 692.79, 692.81, 692.82, 692.83, 692.84, 692.89, 692.9	L24.0, L24.1, L24.2, L25.1, L25.3, L25.4, L25.5, L57.8, L55.0, L55.9, L56.0, L56.1, L56.2, L57.1, L57.5, L57.9, L56.5, L55.1, L55.2, L56.8, L25.0, L58.9, L23.0, L24.81, L23.81, L25.2, L25.8, L25.9
Other and unspecified disorders of back	724.00, 724.01, 724.02, 724.03, 724.09, 724.1, 724.2, 724.3, 724.4, 724.5, 724.6, 724.70, 724.71, 724.79, 724.8, 724.9	M48.00, M48.04, M48.06, M48.08, M54.6, M54.5, M54.30, M54.14, M54.15, M54.16, M54.17, M54.89, M54.9, M43.27, M43.28, M53.2X7, M53.3, M53.2X8, M54.08, M43.8X9, M53.9

Table A4. BMI cutoff values

For those aged 20 years or younger, we used the following criteria based on the CDC's recommended youth BMI-for-age cutoffs:
• **Underweight:** <5th percentile
• **Normal:** 5th to <85th percentile
• **Overweight:** 85th to <95th percentile
• **Obese:** ≥ 95th percentile
• Unknown: missing height and/or weight
For those aged 21 years or older, or when age was missing, we used the following criteria based on the CDC's recommended cutoffs for adults:
• **Underweight:** BMI <18.5
• **Normal:** BMI ∈ [18.5, 25)
• **Overweight:** BMI ∈ [25, 30)
• **Obese:** BMI ≥ 30
• Unknown: missing height and/or weight

Table A5. Calculation of Community Socioeconomic Deprivation Index

An index of socioeconomic deprivation incorporating six indicators from the 2015-2019 American Community Survey 5-year estimates from the US Census:
• Percent less than high school education
• Percent in poverty
• Percent not in the labor force
• Percent on public assistance
• Percent does not own a vehicle
• Percent civilian unemployment
The six indicators were standardized for direction, natural log-transformed, if necessary, z-scored using the standard deviations for Pennsylvania, and summed to create the final, unitless index for each county, township, or census tract. The total number of communities was divided into quartiles of socioeconomic deprivation index. Higher values of the index reflect greater community socioeconomic deprivation.

B. Detailed Cohort Characteristics

Patient level

Patient-level covariates included: race, ethnicity, sex, family history of asthma, and diagnosis of type II diabetes. Race, sex, and family history of asthma were time-invariant. Diagnoses of type II diabetes was time-varying in that they did not have the condition prior to their first diagnosis. Counts shown below for these variables are based on the total cohort of n=46,676 patients.

Race was self-reported in the EHR and was categorized from 19 options to the following, shown in the table below. Approximately 85% of patients identified as white and 12% identified as black.

Patients by collapsed race category

Race category	Number	Percent
White	39,621	84.9
Black	5,524	11.8
Unknown	894	1.9
All other races	637	1.4

Ethnicity was self-reported in the EHR and categorized as shown in the table below. Due to the very small proportion of Hispanic patients, this covariate was not included in the models.

Patients by collapsed ethnicity category

Ethnicity category	Number	Percent
Not Hispanic	44,414	95.2
Unknown	1,887	4.0
Hispanic	375	0.8

Self-reported sex was available from EHR. Nearly 60% of patients were female.

Patients by sex provided in EHR

Sex	Number	Percent
Female	27,337	58.6
Male	19,339	41.4

There were n = 7,209 patients (15.4%) with a family history of asthma. Most had a history of asthma in their biological mother only or biological father only.

Patients by family history of asthma including offspring as first-degree relatives

Family history of asthma	Number	Percent
No	39,467	84.6
Yes	7,209	15.4

About 8% of the cohort has at least one primary diagnosis for type II diabetes.

Patients by primary type II diabetes diagnosis

Type II diabetes	Number	Percent
No	42,865	91.8
Yes	3,811	8.2

Event-level

Visits by year are shown below. There were a higher proportion of visits in 2014 and 2015 and a lower proportion in 2020, which could be indicative of incomplete ascertainment for that year.

Year	Number	Percent
2011	3,274	8.06
2012	3,911	9.63
2013	4,324	10.64
2014	5,149	12.67
2015	5,106	12.57
2016	4,253	10.47
2017	4,158	10.23
2018	4,002	9.85
2019	4,076	10.03
2020	2,377	5.85

Season is shown below. Summer had the fewest number of events.

Season	Number	Percent
Fall	11,425	28.12
Spring	9,937	24.46
Summer	8,506	20.94
Winter	10,762	26.49

The frequency and percent by age group are shown below. Ages 5-13 had the greatest number of events while ages 75-90 had the fewest number.

Age Group	Number	Percent
[5, 13)	11,105	27.33
[13, 19)	3,974	9.78
[19, 45)	10,326	25.41
[45, 62)	9,141	22.50
[62, 75)	4,677	11.51
[75, 90]	1,407	3.46

Information on BMI is shown below. Less than 3% were missing BMI; 42% were obese and 33% were not overweight or obese.

BMI Category	Number	Percent
Not Overweight or Obese	13,341	32.84
Overweight	9,148	22.52
Obese	16,957	41.74
Missing	1,184	2.91

Smoking status is shown in the table below. The majority of events were associated with never smokers, while 9% of events had missing smoking information.

Smoking Status	Number	Percent
Never smoker	25,353	62.40
Current smoker	4,426	10.89
Former smoker	7,134	17.56
Unknown/missing data	3,717	9.15

There were 509 communities represented among the participants. The communities were divided into quartiles to form the cut points (approximately 127 communities in each quartile). Community-level socioeconomic deprivation index by quartile is shown below. Over half of the events were in communities in the highest (best) quartile; 16% were in the lowest (4^{th} quartile).

SES Quartile	Number	Percent
Q1	20,875	51.38
Q2	7,851	19.32
Q3	5,297	13.04
Q4	6,607	16.26

References

1. Gallegos TJ, Varela BA. Trends in hydraulic fracturing distributions and treatment fluids, additives, proppants, and water volumes applied to wells drilled in the United States from 1947 through 2010: data analysis and comparison to the literature. Scientific Investigations Report. 2014-5131. Reston, VA: 2015. (http://pubs.er.usgs.gov/publication/sir20145131).
2. Clark C, Burnham A, Harto C, Horner R. Hydraulic fracturing and shale gas production: technology, impacts, and policy. 2012.
3. McKenzie LM, Witter RZ, Newman LS, Adgate JL. Human health risk assessment of air emissions from development of unconventional natural gas resources. Science of The Total Environment 2012;424:79-87. DOI: https://doi.org/10.1016/j.scitotenv.2012.02.018.
4. Adgate JL, Goldstein BD, McKenzie LM. Potential Public Health Hazards, Exposures and Health Effects from Unconventional Natural Gas Development. Environmental Science & Technology 2014;48(15):8307-8320. DOI: 10.1021/es404621d.
5. Stacy SL, Brink LL, Larkin JC, et al. Perinatal Outcomes and Unconventional Natural Gas Operations in Southwest Pennsylvania. PLOS ONE 2015;10(6):e0126425. DOI: 10.1371/journal.pone.0126425.
6. Casey JA, Savitz DA, Rasmussen SG, et al. Unconventional natural gas development and birth outcomes in Pennsylvania, USA. Epidemiology 2016;27(2):163-72. (In eng). DOI: 10.1097/ede.0000000000000387.
7. Ma Z-Q. Time Series Evaluation of Birth Defects in Areas with and without Unconventional Natural Gas Development. Journal of Epidemiology and Public Health Reviews 2016;1. DOI: 10.16966/2471-8211.107.
8. Rasmussen SG, Ogburn EL, McCormack M, et al. Association Between Unconventional Natural Gas Development in the Marcellus Shale and Asthma Exacerbations. JAMA Internal Medicine 2016;176(9):1334-1343. DOI: 10.1001/jamainternmed.2016.2436.
9. Stacy S. A Review of the Human Health Impacts of Unconventional Natural Gas Development. Current Epidemiol Review 2017;4(1). DOI: doi.org/10.1007/s40471-017-0097-9.
10. Whitworth KW, Marshall AK, Symanski E. Maternal residential proximity to unconventional gas development and perinatal outcomes among a diverse urban population in Texas. PLoS One 2017;12(7):e0180966. (In eng). DOI: 10.1371/journal.pone.0180966.
11. Tustin Aaron W, Hirsch Annemarie G, Rasmussen Sara G, Casey Joan A, Bandeen-Roche K, Schwartz Brian S. Associations between Unconventional Natural Gas Development and Nasal and Sinus, Migraine Headache, and Fatigue Symptoms in Pennsylvania. Environmental Health Perspectives 2017;125(2):189-197. DOI: 10.1289/EHP281.
12. Walker Whitworth K, Kaye Marshall A, Symanski E. Drilling and Production Activity Related to Unconventional Gas Development and Severity of Preterm Birth. Environ Health Perspect 2018;126(3):037006. (In eng). DOI: 10.1289/ehp2622.
13. Koehler K, Ellis JH, Casey JA, et al. Exposure Assessment Using Secondary Data Sources in Unconventional Natural Gas Development and Health Studies. Environmental Science & Technology 2018;52(10):6061-6069. DOI: 10.1021/acs.est.8b00507.

14. Lewis C, Greiner LH, Brown DR. Setback distances for unconventional oil and gas development: Delphi study results. PLOS ONE 2018;13(8):e0202462. DOI: 10.1371/journal.pone.0202462.
15. Casey JA, Goin DE, Rudolph KE, et al. Unconventional natural gas development and adverse birth outcomes in Pennsylvania: The potential mediating role of antenatal anxiety and depression. Environ Res 2019;177:108598. (In eng). DOI: 10.1016/j.envres.2019.108598.
16. Deziel NC, Brokovich E, Grotto I, et al. Unconventional oil and gas development and health outcomes: A scoping review of the epidemiological research. Environmental Research 2020;182:109124. DOI: https://doi.org/10.1016/j.envres.2020.109124.
17. Gonzalez DJX, Sherris AR, Yang W, et al. Oil and gas production and spontaneous preterm birth in the San Joaquin Valley, CA: A case-control study. Environ Epidemiol 2020;4(4):e099. (In eng). DOI: 10.1097/EE9.0000000000000099.
18. Willis MD, Hill EL, Boslett A, Kile ML, Carozza SE, Hystad P. Associations between Residential Proximity to Oil and Gas Drilling and Term Birth Weight and Small-for-Gestational-Age Infants in Texas: A Difference-in-Differences Analysis. Environ Health Perspect 2021;129(7):77002. (In eng). DOI: 10.1289/EHP7678.
19. Deziel NC, Clark CJ, Casey JA, Bell ML, Plata DL, Saiers JE. Assessing Exposure to Unconventional Oil and Gas Development: Strengths, Challenges, and Implications for Epidemiologic Research. Current Environmental Health Reports 2022;9(3):436-450. DOI: 10.1007/s40572-022-00358-4.
20. Finkel ML, Hays J. Environmental and health impacts of 'fracking': why epidemiological studies are necessary. J Epidemiol Community Health 2016;70(3):221-2. (In eng). DOI: 10.1136/jech-2015-205487.
21. Wilke RA, Freeman JW. Potential Health Implications Related to Fracking. JAMA 2017;318(17):1645-1646. (In eng). DOI: 10.1001/jama.2017.14239.
22. Currie J, Greenstone M, Meckel K. Hydraulic fracturing and infant health: New evidence from Pennsylvania. Science Advances 2017;3(12):e1603021. DOI: 10.1126/sciadv.1603021.
23. McKenzie LM, Guo R, Witter RZ, Savitz DA, Newman LS, Adgate JL. Birth Outcomes and Maternal Residential Proximity to Natural Gas Development in Rural Colorado. Environ Health Perspect2014:412-7.
24. McKenzie LM, Allshouse W, Daniels S. Congenital heart defects and intensity of oil and gas well site activities in early pregnancy. Environment International 2019;132:104949. DOI: https://doi.org/10.1016/j.envint.2019.104949.
25. The Process of Unconventional Natural Gas Production. EPA.
26. Webb E, Hays J, Dyrszka L, et al. Potential hazards of air pollutant emissions from unconventional oil and natural gas operations on the respiratory health of children and infants. Rev Environ Health 2016;31(2):225-43. (In eng). DOI: 10.1515/reveh-2014-0070.
27. Willis MD, Jusko TA, Halterman JS, Hill EL. Unconventional natural gas development and pediatric asthma hospitalizations in Pennsylvania. Environ Res 2018;166:402-408. (In eng). DOI: 10.1016/j.envres.2018.06.022.
28. Holguin F, Cardet JC, Chung KF, et al. Management of Severe Asthma: a European Respiratory Society/American Thoracic Society Guideline. European Respiratory Journal 2019:1900588. DOI: 10.1183/13993003.00588-2019.

29. Nuttall FQ. Body Mass Index: Obesity, BMI, and Health: A Critical Review. Nutrition today 2015;50(3):117-128. (In eng). DOI: 10.1097/NT.0000000000000092.