

## Article

# The effects of hydraulic fracturing activities on birth outcomes are evident in a non-individualized county-wide aggregate data sample from Colorado

Chrystal L. Erickson,<sup>1</sup> Ileana G. Barron,<sup>2</sup> Isain Zapata<sup>1</sup>

<sup>1</sup>Department of Biomedical Sciences, Rocky Vista University, Parker, CO; <sup>2</sup>School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA

## Abstract

**Background:** There is growing concern about the recent increase in oil and gas development using hydraulic fracturing. Studies linking adverse birth outcomes and maternal proximity to hydraulic fracturing wells exist but tend to use individualized maternal and infant data contained in protected health care records. In this study, we extended the findings of these past studies to evaluate if analogous effects detected with individualized data could be detected from non-individualized county-wide aggregated data.

**Design and methods:** This study used a retrospective cohort of 252,502 birth records from 1999 to 2019 gathered from a subset sample of 5 counties in the state of Colorado where hydraulic fracturing activities were conducted. We used Generalized Linear Models to evaluate the effect of county-wide well density and production data over unidentified birth weight, and prematurity data. Covariates used in the model were county-wide statistics sourced from the US Census.

**Results:** Our modeling approach showed an interesting effect where hydraulic fracturing exposure metrics have a mixed effect directional response. This effect was detected on birth weight when well density, production and their interaction are accounted for. The interaction effect provides an additional interpretation to discrepancies reported previously in the literature. Our approach only detected a positive association to prematurity with increased production.

**Conclusions:** Our findings demonstrate two main points: First, the effect of hydraulic fracturing is detectable by using county-wide unidentified data. Second, the effect of hydraulic fracturing can be complicated by the number of operations and the intensity of the activities in the area.

## Introduction

Hydraulic fracturing is a highly contentious topic. There is great fear and anxiety over the implications of this unconventional oil drilling practice on air and groundwater pollution. In Colorado,

this concern that has led to several recent local and statewide ballot measures to either restrict wells a certain distance from homes and schools, or to stop this practice all together.<sup>1,2</sup> Hydraulic fracturing involves drilling vertically then horizontally for several miles to access shale embedded with natural gas or coalbed methane. During the process, a pressurized mixture of sand, water, and proprietary fracking fluid is injected into wellbores, fracturing the rock and unlocking trapped hydrocarbons.<sup>3</sup> Many of these chemicals used as hydraulic fracturing fluids are known carcinogens and have been linked to reproductive or developmental toxicity.<sup>4,5</sup> According to the U.S. Geological Survey (USGS), since the end of the 20th century, the use of this technique to produce oil and gas from previously unproductive formations has dramatically increased, which has pushed hydraulic fracturing and related processes into regions where oil and gas had not previously been produced.<sup>6</sup> The evolution and prevalence of this drilling practice is worldwide as an alternative to coal mining for energy production.<sup>7</sup>

With this relatively new proliferation of unconventional drilling, only recently have studies begun to emerge evaluating evidence of long-term negative health effects of hydraulic fracturing.<sup>4,5,8-11</sup> This research has shown there are negative environmental and health impacts for those residing in areas where hydraulic fracturing takes place. These negative effects include higher illness rates, negative birth outcomes, and other changes to morbidity and mortality when compared to areas where there is no fracking.<sup>4,5,12</sup> The only report evaluating such effect in Colorado found positive associations between density and proximity to natural gas wells within a 10-mile radius of maternal residence and prevalence of congenital heart disease and possibly neural tube defects.<sup>9</sup> However, several other studies evaluating the effects of hydraulic fracturing on birth outcomes, including, birth weight in other states and Canada; have had mixed results.<sup>12,13</sup> The majority of studies have reported a negative association that varies in effect size,<sup>14-17</sup> but contradictory results that include positive associations with certain caveats have also been reported in addition to Colorado.<sup>18,19</sup> Birth weight of an infant is an important determinant of its chances of survival and healthy growth and development.<sup>20</sup> Because birth weight is conditioned by the health and

### Significance for public health

*The proliferation of hydraulic fracturing for oil and natural gas production has led to an increase in interest in the public health impact of this industry. Research in this field can be complicated due to data accessibility and concerns of privacy violations. In this study we focus on the assessment of maternal health outcomes while considering data privacy. The main goal of our study was to evaluate the potential of using non-individualized, county-wide data to detect the effects of hydraulic fracturing activities on birth outcomes. This goal was achieved by using county-wide exposure metrics of hydraulic fracturing well density and production and by adjusting to known demographic covariates sourced from Census data. Our study provides an alternate approach to evaluate health effects of hydraulic fracturing activities and provides additional evidence highlighting the complicated effects associations that should be considered in further studies.*

nutritional status of the mother, the proportion of infants born with low birth weight closely reflects the health status of the communities into which they are born and thus is used as a health indicator by the CDC to assess the health of the nation.<sup>21</sup> A normal term birth weight in the United States is between 2,500 and 4,000 grams.<sup>22</sup> It is unclear how evidence of association between birthweight and hydraulic fracturing can still be inconclusive when recent cohort studies have included very large sample sizes of over a million<sup>15</sup> and close to 3 million<sup>18</sup> births along with very accurate covariate data to adjust for confounding variables. In contrast, prematurity has been more consistently associated to hydraulic fracturing activities where a higher exposure is associated with increased risk of preterm birth<sup>12,18</sup> with only a single report showing no association.<sup>17</sup>

A common strategical pattern in these studies is the use of well proximity to the mother's site of residence as a proxy estimator of exposure. This proxy estimator approach has provided very important evidence to associate hydraulic fracturing activities to health-related outcomes; however, the use of such strategy requires birth location data along with the mother's and infant's clinical data, which can potentially identify cohort subjects causing privacy issues. Privacy issues are at this moment a very strong public concern that require urgent real and equitable solutions.<sup>23</sup> Unfortunately, public concerns about the handling of protected data reduces a research team's access to valuable data.<sup>24,25</sup> For this reason, and in view of the healthcare burden and public health concern for hydraulic fracturing activities, we took a similar approach to previous studies but only using non-identifiable county-wide data to explore the wide utility of the methodology. The use of non-identifiable data can reduce the burden for researchers, facilitating and accelerating discovery in the field while addressing privacy concerns. Therefore, the objective of this study was to evaluate generalized effects of hydraulic fracturing exposure from aggregated birth weight and gestational age data (measured as increased risk of additional weeks of premature birth) collected from a subset sample of 5 counties across the state of Colorado. The approach presented in this study opens up the rationale for creative research approaches that can facilitate research advancement in a way that does not invade the privacy of individuals.

## Design and methods

### Experimental design

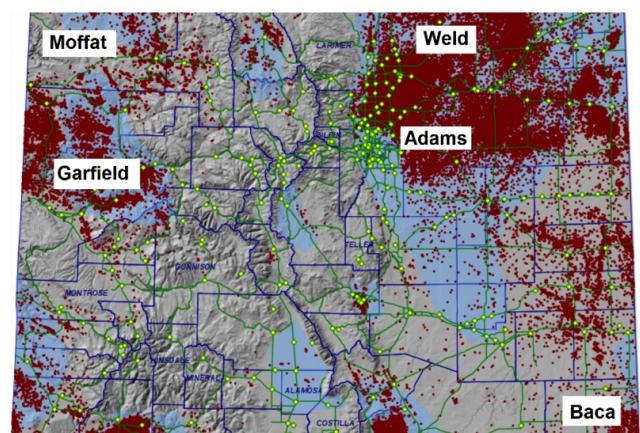
Our study was designed to evaluate the utility of non-identifiable county-wide birth weight and prematurity records with generalized production metrics as proxies for hydraulic fracturing exposure data. This evaluation was performed while adjusting for confounding demographic variables sourced from generalized census data by county. Since our study was designed with the goal of being able to detect birth weight and prematurity effects based on county-wide data, which could be expected to be noisier in comparison to individualized data sets, only a small subset of counties in the state were included. The five Colorado counties included in the study, Adams, Baca, Garfield, Moffat, and Weld, covered a wide range of characteristics in terms of population metrics and hydraulic fracturing activities. Counties like Adams and Weld counties are semi-urban and densely populated while counties like Baca are rural and very sparsely populated. The geographical location of these counties is presented in Figure 1. Our study was vetted by the Institutional Review Board (IRB #2020-0023).

### Birth outcomes data

Although the information is publicly available, a formal request of data was submitted to the Colorado Department of Public Health and Environment's Vital Birth Statistics registry. Specifically, annual, county-wide birth statistics for the five Colorado counties included birth weight (in grams), estimated gestational age at birth (in weeks) and the sex of the infant. Each data set included all babies born between 1999 and 2019 calendar years. No additional statistics or personal data associated with the mothers or infants was attached to this data. The initial dataset included 277,837 births; incomplete records and cases of extremely prematurity (<28 weeks) were excluded from the study. The exclusion of extreme premature babies from the study was used to avoid adding into the data any bias associated to survivability of these babies. This is because survival of babies born at an extremely premature term are highly dependent on access to a neonatal intensive care unit which is a variable not considered in the study,<sup>26</sup> which can also be associated to complex demographic factors.<sup>27</sup> The final data set included 252,502 births (90.88% of the initial data set). For the purpose of adjusting birth weight models, we used preterm birth as a covariate which is defined as a live birth delivered before 37 completed weeks of gestation. However, in our models we used prematurity defined as the preterm weeks by subtracting the estimated gestational age from the maximum value in the data set (45 weeks). This conversion helped the estimation by inverting the scale direction (higher value = higher prematurity), focusing the range to what is relevant for interpretation purposes (as the viability of a birth is greatly reduced with increased prematurity).

### Exposure

Hydraulic fracturing well data was collected through the Colorado Oil and Gas Conservation Commission's website (available at: <https://cogcc.state.co.us/data.html#/cogis>). This information is publicly available. Data collected included: Number of wells in active production and total gas production by county reported yearly. Data was matched by both county and year for the same time range of 1999 to 2019 as in the birth outcomes dataset. Well density was calculated by year per county by dividing the number of wells by the total surface area of the county. Production values were log transformed prior to analysis to reduce scale issues.



**Figure 1.** Location of the five Colorado counties included in this study. Red dots indicate the location of hydraulic fracture wells. Counties in the study are labeled in the map. The original map was developed by the Colorado Department of Natural Resources, Oil & Gas Conservation Commission.

### Covariates

Confounding demographic data that has been previously associated to birth weight and prematurity risk were compiled from the 2000 and 2010 US Census reports and from the 2019 American Community Survey (ACS) (data available at <https://data.census.gov/cedsci/>). All covariate data is publicly available. Data gathered included the following categories: *Population*, which included total county population and population density (calculated in the same manner as well density, total population divided by total surface area of the county); *Age*, which included the percentage of population under 5 years of age, percentage of population under 18 years of age and percentage of population over 65 years of age; *Gender* which only include the female percentage ratio; *Race*, which included the percentage of households that identify as Caucasians (alone), African-Americans, Asians, Native Americans in addition to Hispanics (although Hispanic is not a race but an ethnicity); *Education*, which included the percentage of the population that completed high school and the percentage of the population that completed a Bachelor’s level degree or above; *Income*, which included adjusted household income (income by household adjusted to 2019 dollars to address inflation) and percentage of population at poverty level (defined by Federal guidelines). Covariates were incorporated into the birth outcomes and exposure data by matching it to the closest census year in the following way: births occurring between 1999 to 2004 were matched to 2000 US Census data; 2005 to 2014 were matched to 2010 Census data; and 2015 to 2019 to 2019 ACS data. All covariate data used in the study is presented as Supplementary Table 1.

### Statistical analysis

Data was analyzed using Generalized Linear Models. Dependent variables were birth weight (in grams) and prematurity (as risk for additional preterm weeks). To reduce collinearity issues by category, each covariate was assessed by including it as a single independent factor in the model and selecting the best performing p-value (smallest value within the category). Variables selected by this single independent factor in the model approach were the same for both models. The birth weight model was set as Gaussian distributed since the original units for this variable are continuous and normally distributed; the prematurity model was set as Poisson distributed since the original values are discrete counts with a Poisson distribution. Preliminary models for birth weight evaluated the additive interaction effects of *Well density\*Production* and *Sex\*Term* (*Term* defined as preterm if the estimated gestational age was below 37 weeks otherwise defined as normal). Significant values for the interaction terms in these preliminary models justified

their inclusion in the final model ( $P < 0.05$ ). For birth weight, the final model estimate is expressed in the original units (grams) and defined as follows:

$$\text{Birth weight (grams)} = \text{intercept} + \text{Well density} | \text{Production} + \text{Sex} | \text{Term} + \text{population} + \text{Age} + \text{Gender} + \text{Race} + \text{Education} + \text{Income} + \text{error}$$

In a similar manner, the prematurity final model was defined through preliminary models where the inclusion of the additive interaction effect of *Well density\*Production* was evaluated. The *Sex\*Term* interaction was not evaluated for prematurity since the variable *Term* duplicates the dependent variable. For prematurity, the final model estimate is expressed in hazards ratio of one-week increase and defined as follows:

$$\text{Prematurity (hazards ratio of one-week increase)} = \text{Well density} | \text{Production} + \text{Sex} + \text{population} + \text{Age} + \text{Gender} + \text{Race} + \text{Education} + \text{Income} + \text{error}$$

All modeling analyses and descriptive statistics were performed in SAS v9.4 (SAS institute, Cary NC) though PROC MEANS and PROC GLIMMIX. Significant differences were declared at an  $\alpha$  threshold of 0.05.

### Results

Our study used a total of 252,502 birth records spanning from 1999 to 2019 in five counties, Adams, Baca, Garfield Moffat and Weld, that allow for hydraulic fracturing activities in the state of Colorado. These counties represent a wide range of demographic characteristics that occur in the state of Colorado. These demographic characteristics span all categories measured in the study, which are presented in detail in Supplementary Table 1. Birth outcomes are presented in Table 1. We observed mean birth weights to significantly vary in some counties ( $p=3.35E-126$ ), along with mean estimated gestational age at birth which also significantly varied in some counties ( $p=9.91E-32$ ), no significant variation of sex ratios by county was observed ( $p=0.1104$ ). However, the observed female percent ratios were all significantly biased towards a lower female proportion and deviated from the expected 50% ( $p$ -value range for Binomial probability of 0.0217 for Baca County to 7.86E-19 for Adams County) Last, preterm ratios did show significant differences by county ( $p=5.20E-8$ ). All these differences are reported although they are not in the scope of our study to evaluate.

**Table 1. Descriptive statistics of birth outcomes. Parameters are described for the full dataset and each of the five counties included in the study.**

| County  | Full dataset     | Adams           | Baca            | Garfield        | Moffat          | Weld            |
|---|------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| <b>Cohort size (n)</b>                            | 252,502          | 150,335         | 836             | 17,334          | 3,902           | 80,095          |
| <b>Birth weight (grams)</b>                       | 3,247.1 ± 506.3  | 3,230.7 ± 505.5 | 3,224.0 ± 521.2 | 3,246.9 ± 489.6 | 3,182.6 ± 491.8 | 3,281.3 ± 509.9 |
| <b>Estimated gestational age at birth (weeks)</b> | 38.8 ± 1.6       | 38.8 ± 1.6      | 38.6 ± 1.7      | 38.8 ± 1.6      | 38.8 ± 1.7      | 38.8 ± 1.6      |
| <b>Sex (Females (% of total))</b>                 | 123,296 (48.83%) | 73,534 (48.91%) | 408 (48.80%)    | 8,473 (48.88%)  | 1,823 (46.72%)  | 39,058 (48.76%) |
| <b>Term (preterm (% of total))</b>                | 19,754 (7.82%)   | 11,934 (7.94%)  | 84 (10.05%)     | 1,266 (7.30%)   | 379 (9.71%)     | 6,091 (7.60%)   |

Exposure to hydraulic fracturing activities was evaluated in our study through the evaluation of production and well density by county per year. Yearly trends for these two-exposure metrics are presented in Figure 2. The production output of these 5 counties varied through the years where mean values were within 3 orders of magnitude from the least producing Baca County with an average 1,710,339 units per year to Garfield County with an average 416,260,300 units per year. Similarly, well density varied within a range of 0.0618 to 4.7536 active wells per square mile as yearly average. Among our five counties evaluated we observed mixed increases and decreases in production and well density through the span of the studied timeframe. It is noteworthy to mention that well density does not directly explain production as the most productive county (Garfield) has 45% lower average well density than the densest county (Weld).

Single factor Generalized Linear Models were used for a preliminary evaluation of covariate inclusion for birth weight and prematurity. These covariates were always significant in these models (*data not shown*) although some categories became non-significant as they were included in the full model along with all other covariate categories. For the birth weight model (Table 2), we successfully detected a strong positive association to exposure variables (well density and production) using county-wide, non-individualized data, which achieves part of our main goal. This association possess a very interesting situation since the well density by production interaction was significant but contributed to the model with a negative parameter estimate. The mix of positive and negative contributions of the two exposure variables is suggestive of a complex association to exposure where not only the proximity, but the amount of activity is important for estimating health effects. Although two recent studies have incorporated production output into their evaluations,<sup>18,28</sup> interaction effects have not been explicitly reported besides main effects. Only one of such studies by van Tran in 2020<sup>18</sup> evaluated birth weight, finding small mixed associations when comparing rural to urban locations while also accounting for active and inactive wells. In the literature, the most common exposure metric used to evaluate the effect of hydraulic fracturing activities is well proximity to the mother's residence, which is determined by geographical location. In addition to the previous

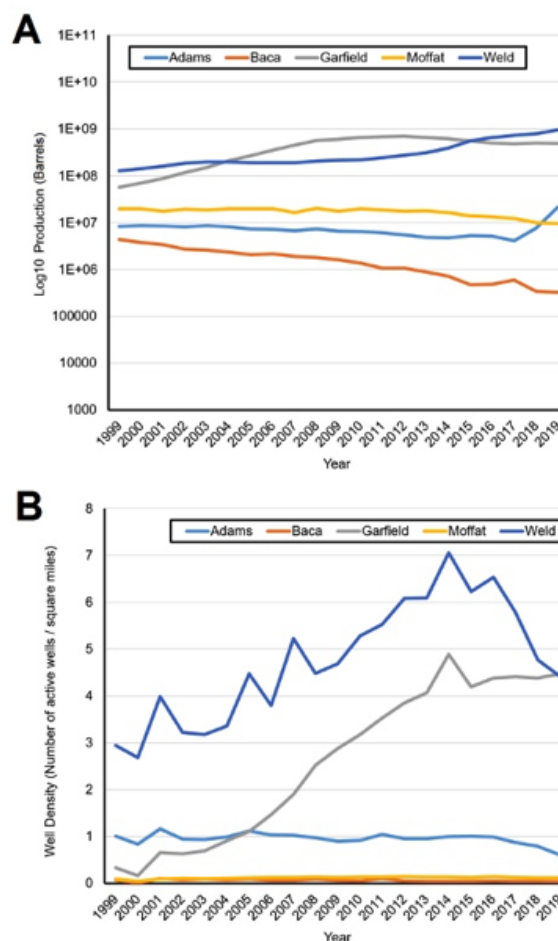


Figure 2. Hydraulic fracturing exposure metrics evaluated in the study. A) Production graph across the five counties evaluated by year from 1999 until 2019, on a logarithmic scale. B) Well density graph across the five counties evaluated by year from 1999 until 2019.

Table 2. Birth weight model results. Fixed effect outcomes and parameter estimates.

|            | Effect                                | P        | Parameter Estimates (grams)  | SE                    |
|------------|---------------------------------------|----------|--|-----------------------|
| Exposure   | Intercept                             | -        | 1641.6200  | 172.37                |
|            | Well density                          | 0*       | 186.9300   | 1.1917                |
|            | Production                            | 4.36E-09 | 14.2897  | 2.4344                |
|            | Well density * Production             | 0*       | -9.2953  | 0*                    |
| Covariates | Sex                                   | 1.7E-200 | M 124.8800<br>F 0*   | 1.8628<br>-           |
|            | Term                                  | 0*       | Pre term -812.5200<br>Normal 0*  | 4.9226<br>-           |
|            | Sex * Term                            | 1.0E-12  | M, Pre term -47.6380<br>M, Normal 0*<br>F, Pre term 0*<br>F, Normal 0* | 6.6843<br>-<br>-<br>- |
|            | Population (Population density)       | 0.9904   | 0.0007   | 0.0598                |
|            | Age (% of Population over 65)         | 0.8642   | 0.2231   | 1.3039                |
|            | Gender (Female % ratio)               | 1.7E-10  | 25.7631  | 4.0349                |
|            | Race (Caucasian alone)                | 0.1301   | 0.9034   | 0.5968                |
|            | Education (Bachelor's level or above) | 0.3476   | 0.5075   | 0.5404                |
|            | Income (% at poverty level)           | 0.0020   | -1.4711  | 0.4764                |

\*Estimated as zero

studies, other studies have incorporated well density into the equation<sup>9,16,17,19</sup> with some of them reporting mixed results; however, by not incorporating additional metrics to account for the production intensity, no further explanation of the results could be offered. Some covariate categories in the full model were also significant, including sex of the infant, *Term* (denominated as preterm or normal at birth) and gender ratio as the strongest effect variables based on absolute values of parameter estimates. The covariate category of Income that represented the percentage of the population at poverty level was also significant but contributed marginally to the model. For the prematurity model (Table 3), our approach was also successful in detecting significant associations using county-wide, non-individualized data. For this model, we observed a different pattern when compared to the association seen with the birth weight data where only production was significantly associated to prematurity while well density and its interaction with production were not. Although not all variables in the exposure category were significant, parameter estimate effect directions were the same as in birth weight with the main effects positively associated while the interaction effect was negatively associated. Our findings closely resemble previous reports where increased preterm risk is consistently associated with exposure to hydraulic fracturing activities.<sup>9,16,18,19,28</sup> For this model, demographic covariates more consistently remained as statistically significant when included in the

full model. Among those categories that remained significant, sex of the infant was strongly associated while gender ratio showed a modest effect based on the parameter estimates. Population, age, education and income categories were also significantly associated, although their contributions to increased risk were much smaller. In addition to presenting significant associations and parameter estimates for both of the full models in Tables 2 and 3, we present an outcome scenario table based exclusively on exposure (Table 4). This table presents an empirical visualization of exposure on birth weight and prematurity. This table was constructed using parameter estimates presented in Tables 2 and 3 and minimum and maximum values for the full dataset. For well density, the minimum and maximum values were 0.0062573 and 7.0582524 active wells per square mile respectively, while for production, they were 328,894 and 955,469,444 respectively. In this table, it is clear how hydraulic fracturing increases birth weight when well density is in the low end (of ~100 gr difference increase) when comparing lowest versus highest production. This difference is reversed when well density is on the high end (of ~400 gr difference decrease) when comparing the lowest versus the highest production. The negative effect of the interaction between these parameters is the source of this discrepancy. A normal birth term cohort model (parameter estimates not shown) is also presented in Table 4 as a comparison, this cohort presents an analogous scenario. A very similar

**Table 3. Prematurity model results. Type 3 test for fixed effect outcomes and parameter estimates.**

|            | Effect                                | P       | Parameter Estimates<br>(hazards ratio of one additional week) |        | SE     |
|------------|---------------------------------------|---------|---|--------|--------|
| Exposure   | Well density                          | 0.2955  | 0.0241  |        | 0.0231 |
|            | Production                            | 5.8E-07 | 0.0162  |        | 0.0032 |
|            | Well density * Production             | 0.2379  | -0.0014   |        | 0.0012 |
| Covariates | Sex                                   | 6.5E-41 | M   | 2.1632 | 0.2045 |
|            |                                       |         | F   | 2.1495 | 0.2045 |
|            | Population (Population density)       | 0.0014  | 0.0002  |        | 0.0001 |
|            | Age (% of Population over 65)         | 1.0E-08 | 0.0070  |        | 0.0012 |
|            | Gender (Female % ratio)               | 0.0007  | -0.0143   |        | 0.0042 |
|            | Race (Caucasian alone)                | 0.8762  | 0.0001  |        | 0.0005 |
|            | Education (Bachelor's level or above) | 8.0E-05 | -0.0026   |        | 0.0007 |
|            | Income (% at poverty level)           | 1.5E-20 | 0.0041  |        | 0.0004 |

**Table 4. Outcome scenarios based exclusively on exposure. Parameter estimates for exposure variables were used to calculate outcome scenarios using the data set's lowest and highest values. This outcome scenario allows for visualization of the effect of well density, production and the interaction effects. For birth weight, all three effects were significant in the full model while for prematurity, production was the only significant effect.**

| Dependent Variable                               | Model                                    | Lowest well density |                    | Highest well density |                    |
|--|--|---------------------|--------------------|----------------------|--------------------|
|  |  | Lowest production   | Highest production | Lowest production    | Highest production |
| Birth weight (grams)                             | Adjusted model                           | 182.0               | 295.4              | 667.5                | 258.2              |
|  | Difference                               | 113.5               |                    | -409.2               |                    |
|  | Normal term babies only (Adjusted model) | 165.9               | 269.4              | 644.1                | 244.7              |
|  | Difference                               | 103.5               |                    | -399.3               |                    |
| Prematurity (Hazards ratio of one week increase) | Adjusted model                           | 0.2057              | 0.3347             | 0.2496               | 0.2994             |
|  | Raw Difference                           | 0.1290              |                    | 0.0497               |                    |
|  | % Difference                             | 13.77               |                    | 5.10                 |                    |

pattern also occurs for prematurity when following the same rationale, although it must be emphasized that for this model only production was significant. Our findings suggest that by evaluating interaction effects in exposure models, researchers can better define complex relationships that often lead to associations with counterintuitive effect directions.

## Discussion

The main goal of our study was to explore the use of non-individualized, county-wide data to estimate the effects of hydraulic fracturing on birth weight and increased prematurity risk. Our study was conceived as the starting point and an alternative to current mainstream methods. The approach we describe can be further refined to incorporate more specialized and detailed metrics, which are likely to yield more precise estimates. For this reason, the intention with this study was not to emphasize directly the parameters estimated but to demonstrate how this approach can be a viable alternate method for analysis. The approach presented here can be useful when privacy concerns and the handling of protected data are likely to be a limitation of future studies.<sup>24,25</sup>

Findings in our study are limited by the representativeness of the sample used. In Colorado, an estimated 74% of counties are classified as rural but only 13% of the total population in the state lives in those counties<sup>29</sup>. For this reason, the counties included in our study were not randomly selected. The five counties sampled were selected for being demographically and geographically similar to population parameters for the whole state. In the sampling considered here, a large proportion of the data (91%) came from Adams and Weld counties, which are two largely suburban and semi-urban counties that are included in the Denver-Aurora Combined Statistical Area. In contrast Baca, Garfield and Moffat counties are largely rural, noting that, by volume, Garfield County is a major producer of natural gas in the state. The diversity of hydraulic fracturing practices along with their demographics and population proportions make the 5 counties chosen a good representation of the state. The representativeness of our sample can be debated and for that reason we avoid specific parameter estimate inference being the focus of our study. Despite the inherent study limitations and challenges posed by using non-identified data, it is noteworthy to mention that this study closely corroborates the findings of the only previous study performed in Colorado by McKenzie *et al.* in 2014<sup>9</sup> where mixed effect associations were detected for birth weight along with a positive association to preterm birth risk using detailed location data along with mother's and the child's information.

The use of non-individualized county-wide aggregated data likely has implications in the signal to noise ratio, which was expected to be larger in our approach. The larger noise that comes from imprecise aggregate data is a welcome challenge for demonstrating the concept we present in this study since being able to detect an effect using lower resolution data dampens the signal and, if detected, implies a strong association. Unfortunately, such premise has also a negative side, because our approach is likely to decrease statistical power, which can be detrimental when precise estimations of weaker effects are required. Additionally, the implementation and validity of our approach is predicated on the premise of a large series of related studies; these studies provide a precedent that is necessary to evaluate and judge the outcomes of this new approach. Due to the limitations discussed previously, this new concept cannot replace traditional mainstream approaches but provides an additional option to explore the data.

## Conclusions

The main goal of our study was to evaluate the potential of using non-individualized, county-wide data to detect the effects of hydraulic fracturing activities on birth outcomes. We achieved this main goal by detecting strong associations between county-wide exposure metrics of well density and production and by adjusting to known demographic covariates that were sourced from Census derived data. More specifically, birth weight was found to be positively associated to well density and production but negatively associated to their interaction effect. This mixed effect direction association to exposure parameters provides an interpretation to the mixed outcomes reported by previous studies. In contrast, we only detected a strong positive association between production and increased prematurity risk, which is concordant with previous studies. In summary, our study provides an alternate approach to evaluate health effects of hydraulic fracturing activities and provides additional evidence highlighting the complicated effects associations that should be considered in further studies.

**Correspondence:** Isain Zapata, Department of Biomedical Sciences, Rocky Vista University, 8401 S. Chambers Road, Parker, CO 80134, USA. Tel. +1.720.874-2493. E-mail: izapata@rvu.edu

**Key words:** Privacy; maternal outcomes; hydraulic fracturing; unidentified.

**Contributions:** CE, contributed to the conception, data acquisition, interpretation of data along with drafting of the work; IB, contributed to the interpretation of data along with drafting of the work; IZ, contributed with conception, design, analysis, interpretation of data along with drafting of the work. All authors approved the final version of the manuscript and agreed to be accountable for all aspects of the work.

**Conflict of interest:** The authors declare no conflicts of interest.

**Acknowledgements:** We would like thank Dr. Amanda Brooks for her assistance in improving the quality of our manuscript.

**Ethics approval and consent to participate:** The study was vetted by the Rocky Vista University Institutional Review Board (IRB #2020-0023). This study used non identifiable data that is available in the public domain.

**Patient consent for publication:** Not applicable.

**Informed consent:** Not applicable.

**Availability of data and materials:** The data used to support the findings of this study are available from the corresponding author upon request.

Received for publication: 25 July 2021.

Accepted for publication: 4 September 2021.

©Copyright: the Author(s), 2021

Licensee PAGEPress, Italy

Journal of Public Health Research 2022;11:2551

doi:10.4081/jphr.2021.2551

This work is licensed under a Creative Commons Attribution NonCommercial 4.0 License (CC BY-NC 4.0).

## References

- Raimi D, Krupnick A, Bazilian M. Partisanship and proximity predict opposition to fracking in Colorado. *Energy Res Soc Sci* 2020;64:101441.
- Minor J. Local Government Fracking Regulation: A Colorado Case Study. *Stanford Environ Law J* 2014;33:59.
- US EPA (Environmental Protection Agency). The process of unconventional natural gas production: hydraulic fracturing. 2013. Accessed: June 30, 2020. Available from: <https://www.epa.gov/uog/process-unconventional-natural-gas-production>
- Bamber AM, Hasanali SH, Nair AS, et al. A Systematic review of the epidemiologic literature assessing health outcomes in populations living near oil and natural gas operations: Study quality and future recommendations. *Int J Environ Res Public Health* 2019;16:2123.
- Elliott EG, Trinh P, Ma X, et al. Unconventional oil and gas development and risk of childhood leukemia: Assessing the evidence. *Sci Total Environ* 2017;576:138-47.
- USGS. When did hydraulic fracturing become such a popular approach to oil and gas production? Accessed: August 11, 2020. Available from: [https://www.usgs.gov/faqs/when-did-hydraulic-fracturing-become-such-a-popular-approach-oil-and-gas-production?qt-news\\_science\\_products=0#qt-news\\_science\\_products](https://www.usgs.gov/faqs/when-did-hydraulic-fracturing-become-such-a-popular-approach-oil-and-gas-production?qt-news_science_products=0#qt-news_science_products)
- Melikoglu M. Shale gas: Analysis of its role in the global energy market. *Renew Sustain Energy Rev* 2014;37:460-8.
- McKenzie LM, Allshouse WB, Byers TE, et al. Childhood hematologic cancer and residential proximity to oil and gas development. *PLoS One* 2017;12:e0170423.
- McKenzie LM, Guo R, Witter RZ, et al. Birth outcomes and maternal residential proximity to natural gas development in rural Colorado. *Environ Health Perspect* 2014;122:412-7.
- McKenzie LM, Crooks J, Peel JL, et al. Relationships between indicators of cardiovascular disease and intensity of oil and natural gas activity in Northeastern Colorado. *Environ Res* 2019;170:56-64.
- McKenzie LM, Allshouse W, Daniels S. Congenital heart defects and intensity of oil and gas well site activities in early pregnancy. *Environ Int* 2019;132:104949.
- Balise VD, Meng C-X, Cornelius-Green JN, et al. Systematic review of the association between oil and natural gas extraction processes and human reproduction. *Fertil Steril* 2016;106:795-819. doi
- Caron-Beaudoin É, Whitworth KW, Bosson-Rieutort D, et al. Density and proximity to hydraulic fracturing wells and birth outcomes in Northeastern British Columbia, Canada. *J Expo Sci Environ Epidemiol* 2021;31:53-61.
- Apergis N, Hayat T, Saeed T. Fracking and infant mortality: fresh evidence from Oklahoma. *Environ Sci Pollut Res* 2019;26:32360-7. d
- Currie J, Greenstone M, Meckel K. Hydraulic fracturing and infant health: New evidence from Pennsylvania. *Sci Adv* 2017;3:e1603021.
- 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:e0180966.
- Stacy SL, Brink LAL, Larkin JC, et al. Perinatal outcomes and unconventional natural gas operations in Southwest Pennsylvania. *PLoS One* 2015;10:e0126425.
- Tran KV, Casey JA, Cushing LJ, Morello-Frosch R. Residential proximity to oil and gas development and birth outcomes in California: A retrospective cohort study of 2006–2015 births. *Environ Health Perspect* 2020;128:67001.
- Casey JA, Savitz DA, Rasmussen SG, et al. Unconventional natural gas development and birth outcomes in Pennsylvania, USA. *Epidemiology* 2016;27:163-72.
- McCormick MC. The contribution of low birth weight to infant mortality and childhood morbidity. *N Engl J Med* 1985;312:82-90.
- United Health Foundation. Natality public-use data. America's Health Rankings analysis of CDC WONDER Online Database. 2021. Accessed: April 7, 2021. Available from: <https://www.healthypeople.gov/2020/leading-health-indicators/2020-LHI-Topics>
- CDC. International Notes Update: Incidence of low birth weight. *MMWR Morb Mortal Wkly Rep* 1984;33:459-460, 465-7.
- Dwork C, Karr AF, Nissim K, Vilhuber L. On privacy in the age of COVID-19. *J Priv Confidentiality* 2020;10:749
- Jeff Smith H, Dinev T, Xu H. Information privacy research: An interdisciplinary review. *MIS Quarterly* 2011;35:989-1015.
- Mallappallil M, Sabu J, Guessner A, Salifu M. A review of big data and medical research. *SAGE Open Med* 2020;8:2050312120934839.
- Patel RM, Rysavy MA, Bell EF, Tyson JE. Survival of infants born at periviable gestational ages. *Clin Perinatol* 2017;44:287-303.
- Sigurdson K, Mitchell B, Liu J, et al. Racial/ethnic disparities in neonatal intensive care: A systematic review. *Pediatrics* 2019;144:e20183114.
- Whitworth KW, Marshall AK, Symanski E. Drilling and production activity related to unconventional gas development and severity of preterm birth. *Environ Health Perspect* 2018;126:037006.
- Colorado Rural Health Center. The State of Health in Rural Colorado; 2016. Available from: <https://www.colorado.gov/pacific/sites/default/files/11 - Colorado Rural Health Center - 2016 Snapshot.pdf>