



Published in final edited form as:

*Epidemiology*. 2016 March ; 27(2): 163–172. doi:10.1097/EDE.0000000000000387.

## Unconventional natural gas development and birth outcomes in Pennsylvania, USA

Joan A. Casey<sup>a,e</sup>, David A. Savitz<sup>d</sup>, Sara G. Rasmussen<sup>a</sup>, Elizabeth L. Ogburn<sup>b</sup>, Jonathan Pollak<sup>a</sup>, Dione G. Mercer<sup>f</sup>, and Brian S. Schwartz<sup>a,c,f</sup>

<sup>a</sup> Department of Environmental Health Sciences, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

<sup>b</sup> Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA

<sup>c</sup> Department of Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA

<sup>d</sup> Department of Epidemiology, Brown University School of Public Health, Providence, Rhode Island, USA and Department of Obstetrics and Gynecology, Alpert School of Medicine, Brown University, Providence, Rhode Island, USA

<sup>e</sup> Robert Wood Johnson Foundation Health and Society Scholars Program, UC San Francisco and UC Berkeley, California, USA

<sup>f</sup> Center for Health Research, Geisinger Health System, Danville, Pennsylvania, USA

### Abstract

**Background**—Unconventional natural gas development has expanded rapidly. In Pennsylvania the number of producing wells increased from zero in 2005 to 3689 in 2013. To our knowledge, no prior publications have focused on unconventional natural gas development and birth outcomes.

**Methods**—We performed a retrospective cohort study using electronic health record data on 9384 mothers linked to 10946 neonates in the Geisinger Health System from January 2009–January 2013. We estimated cumulative exposure to unconventional natural gas development activity with an inverse-distance squared model that incorporated distance to the mother’s home; dates and durations of well pad development, drilling, and hydraulic fracturing; and production volume during the pregnancy. We used multilevel linear and logistic regression models to examine associations between activity index quartile and term birth weight, preterm birth, low 5 minute Apgar score and small size for gestational age, while controlling for potential confounding variables.

**Results**—In adjusted models, there was an association between unconventional natural gas development activity and preterm birth that increased across quartiles, with a fourth quartile odds ratio of 1.4 (95% CI: 1.0-1.9). There were no associations of activity with Apgar score, small for

---

**Corresponding author:** Brian S. Schwartz, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Room W7041, Baltimore, MD, 21205. bschwar1@jhu.edu..

**Conflicts of Interest** We declare that we have no conflicts of interest.

gestational age, or term birth weight (after adjustment for year). In a *post-hoc* analysis, there was an association with physician-recorded high-risk pregnancy identified from the problem list (fourth vs. first quartile, 1.3 [95% CI: 1.1-1.7]).

**Conclusion**—Prenatal residential exposure to unconventional natural gas development activity was associated with two pregnancy outcomes, adding to evidence that unconventional natural gas development may impact health.

---

## INTRODUCTION

The last decade has seen rapid development of unconventional natural gas resources worldwide; the International Energy Agency reports that 18% of global gas production now comes from unconventional sources. The steepest increases have occurred in the United States (U.S.) and in particular in the Marcellus shale in Pennsylvania. From 2006 to 2013, annual conventional gas production in Pennsylvania was stable at around 5.7 billion cubic meters (bcm); prior to 2009, unconventional production was less than 10 bcm, and then production increased rapidly to 3048 bcm in 2013.

Unconventional natural gas development is a large-scale multi-stage process.<sup>1-4</sup> Developers use diesel equipment to clear land for well pads, transport materials, and drill multiple wells per pad. Directional drilling, first vertically and then horizontally, and hydraulic fracturing (“fracking”) differentiate this process from conventional development. Hydraulic fracturing involves injecting millions of liters of water mixed with sand and chemicals into the borehole causing fractures in the shale formation. Fracturing fluids, flowback and produced water, and natural gas then flow to the surface for collection and use. Gas is sometimes flared, releasing pollutants. Wells produce natural gas at high rates for the first year, with a rapid decline over the first three years.

Prior studies have demonstrated environmental impacts from the various stages of unconventional natural gas development including pollution of air,<sup>5-9</sup> surface water,<sup>10</sup> groundwater,<sup>11,12</sup> and soil as recently reviewed.<sup>1-3</sup> Truck traffic, drilling, hydraulic fracturing, and production can generate diesel particulate matter, fine particulate matter (PM<sub>2.5</sub>), methane, NO<sub>x</sub>, and volatile organic compounds, which are also ozone precursors.<sup>5-7,13</sup> Some of these pollutants, most consistently PM<sub>2.5</sub>, NO<sub>x</sub>, SO<sub>x</sub>, and ozone, have been associated with adverse birth outcomes including low or reduced birth weight<sup>14-16</sup> and preterm birth.<sup>14,17,18</sup> PM<sub>2.5</sub> and ozone are regional air pollutants, so women living long distances from unconventional natural gas development could experience effects.

Expectant mothers could also be exposed to water pollution from unconventional natural gas development. A recent study identified 2-n-butoxyethanol – a chemical found in flowback water from the process, which might be a general indicator of its contamination – in household well water in Pennsylvania.<sup>12</sup> In addition, people living in communities near unconventional natural gas development commonly report symptoms (e.g., upper respiratory symptoms, headaches), and may experience psychosocial stressors from rapid industrial development, increased motor vehicle traffic, potential influences on environmental radon pathways, noise, and infusion of short-term workers.<sup>1,4,19-23</sup> Some of these exposures have

also been linked to negative birth outcomes.<sup>24,25</sup> A recent study in Colorado reported that density of and proximity to natural gas wells were associated with congenital heart and neural tube defects, but not with birth weight or preterm birth.<sup>26</sup> This study did not distinguish between conventional and unconventional wells, and mainly described associations with conventional wells since the Energy Information Agency estimated that only 25% of natural gas produced in Colorado in 2009 came from unconventional sources. There is an unpublished study that found mothers living near unconventional natural gas development in Pennsylvania gave birth to infants with increased prevalence of low birth weight, low Apgar scores, and small for gestational age.<sup>27</sup>

In this study, we exploited the geographic overlap of the Geisinger Health System and unconventional natural gas development in Pennsylvania to conduct a retrospective cohort study by linking electronic health record data to estimates of exposure to the activities during pregnancy. Despite calls for health studies,<sup>28,29</sup> to our knowledge there is only one published population-based study focused on unconventional natural gas development and objective health outcomes.<sup>30</sup> We evaluated associations between an index of unconventional natural gas development activity and four birth outcomes.

## METHODS

### Study area and participants

The Geisinger Health System serves a primary market of approximately 40 counties in central and northeast Pennsylvania, a region with a 2010 population of over 4 million residing in over 1200 communities defined as townships, boroughs, and census tracts in cities.<sup>31</sup> Patients with a Geisinger primary care provider are representative of the general population based on age, sex, race/ethnicity, and rural residence.<sup>32</sup> Neonates were delivered at two hospitals, Geisinger Medical Center in Danville, which has a Level IV neonatal intensive care unit (NICU), and Geisinger Wyoming Valley in Wilkes-Barre, which has a Level II NICU. The Institutional Review Board at the Geisinger Health System reviewed and approved the study.

Singleton births to women who delivered at Geisinger between 2006 and January 2013 were eligible for inclusion. We identified births and deliveries using International Classification of Diseases, Ninth Revision codes (i.e., V27.x, V30.x) in mother and neonate electronic health records. We used medical record numbers and other data found in the electronic health record to link mothers with their neonates. We excluded those whom we could not match, stillbirths, and neonates with serious birth defects, birth weights < 500g or gestational ages < 22 weeks. Only mother's 2013 address was available from the electronic health record, so we assumed they lived at the same address during pregnancy. We geocoded women's residences using ArcGIS 10.2<sup>31</sup> and excluded those who did not reside in Pennsylvania or whose address we were unable to geocode. We evaluated our assumption of mother's residential stability by comparing addresses in two Geisinger Health System datasets, 39 months apart (one from 2010 and the other from 2013), among 333,322 patients in both datasets. Due to strong collinearity between the unconventional natural gas development exposure metric and calendar year, we also excluded births prior to 2009 when little such activity had taken place in the study region.

## Birth outcomes

We extracted data from electronic health record files including labor and delivery notes and a separate labor and delivery database maintained continuously by nursing personnel. The clinician recorded gestational age as part of routine care based on patient-reported last menstrual period and 20 week ultrasound. We estimated the first day of pregnancy from gestational age. We studied four birth outcomes: term ( $\geq 37$  week) birth weight, preterm birth ( $< 37$  weeks gestation), low 5 minute Apgar score ( $< 7$ ), and small for gestational age; we isolated moderate to late preterm birth (32-36 weeks gestation) in a sensitivity analysis. Infants with low 5 minute Apgar scores often require respiratory support and have poorer future academic achievement.<sup>33</sup> Small for gestational age was defined as less than the sex-specific 10<sup>th</sup> percentile of weight for each week of gestation within the Geisinger population from 2006-2013. While creating the *a priori* outcomes, we discovered that maternal and fetal specialists often use the electronic health record problem list to identify a pregnancy as high-risk. Because we hypothesized that UNGD could contribute to conditions (e.g., pulmonary, cardiovascular) that could designate a pregnancy as high-risk, *post hoc* we added high-risk pregnancy as an outcome.

## Unconventional natural gas development activity index

We collected data, spanning 2005-2013, on well drilling and production dates and volumes from the Pennsylvania Department of Environmental Protection and on well stimulation dates and drilling depth from the Pennsylvania Department of Conservation and Natural Resources. We collaborated with SkyTruth (Shepherdstown, WV, skytruth.org) to use crowdsourcing to confirm well pad locations using U.S. Department of Agriculture aerial photographs. We imputed missing total depths, production volumes, and stimulation dates from available data. The assembled dataset included latitude and longitude of each well; dates of well spudding (i.e., beginning of drilling), perforation, stimulation, and production; total well depth; volume of natural gas produced; and the number of producing days annually. Because phases of unconventional natural gas development (i.e., pad development, drilling, stimulation, production) are known to differ by exposures and duration, we derived individual-level estimates to each of these four phases. Although there was heterogeneity by well, for the purposes of exposure assignment, we used published descriptions<sup>34</sup> of the process and information in our own data to estimate phase durations: (1) pad development = the 30 days prior to the first well drilled on a pad; (2) drilling = 1-30 days, based on total well depth; (3) hydraulic fracturing = 7 days; and (4) production = present when reported production values were non-zero.

We first created four exposure metrics by phase that incorporated all wells statewide as:

$$Mother\ j\ metric = \sum_{i=1}^n \sum_{k=1}^l m(I_A(k)) / d_{ij}^2$$

where  $n$  was the number pads or wells;  $k$  was the day with 1 equal to January 1, 2009 and  $l$  was equal to 1125 or January 31, 2013;  $m$  was 1 for pad and drilling,  $m$  was total well depth for stimulation (because we used total well depth as a surrogate for truck trips and hydraulic

fracturing fluid volume), and  $m$  was gas volume for production (because we used production volume as a surrogate for air pollution emissions);  $I_A(k)$  was 1 when the phase overlapped temporally with gestation; and  $d_{ij}^2$  was the squared-distance between the coordinates of pad or well  $i$  and mother  $j$ 's home address. The phase-specific units were pads/m<sup>2</sup>, wells/m<sup>2</sup>, total well depth (m)/m<sup>2</sup>, and gas production volume m<sup>3</sup>/m<sup>2</sup> for pad, drilling, stimulation, and production metrics, respectively. The denominator was always the squared-distance between wells and residences (m<sup>2</sup>).

Because we wanted to estimate exposure to phases of unconventional natural gas development and there was collinearity between the four exposure metrics ( $\rho$ , 0.6-0.9), each was z-transformed then summed to estimate the unconventional natural gas development activity index (hereafter referred to as the activity index). This meant that a woman living close to several well pads under development, but far from any producing wells could have a similar index as a woman living near only producing wells. We did not evaluate trimester-specific indices because of very high inter-trimester correlations. We divided the aggregated activity index into quartiles for analysis.

### Covariates

We included clinical, demographic, and environmental covariates to control for potential confounding based on *a priori* hypotheses and previous studies of birth outcome risk factors including neonate sex, gestational age (for birth weight), season and year of birth, maternal age, race/ethnicity, Geisinger primary care provider status, smoking status during pregnancy, pre-pregnancy body-mass index (BMI), parity, antibiotic orders during pregnancy, and receipt of Medical Assistance, a surrogate for low family socioeconomic status.<sup>35,36</sup> For teenagers (< 20 years), we categorized pre-pregnancy BMI using z-scores based on U.S. Centers for Disease Control and Prevention data. Environmental covariates included distance to nearest major road (principal arterial and larger based on U.S. Census Bureau Topologically Integrated Geographic Encoding and Referencing road files),<sup>24,37</sup> community socioeconomic deprivation<sup>38</sup> and residential greenness (based on the average normalized difference vegetation index values in the 1250m × 1250m area surrounding the residence in the three seasons prior to delivery).<sup>39</sup> Due to concern about the potential contamination of ground water in the region, we used Pennsylvania Department of Environmental Protection public water service areas to assign household water source as municipal or well water.<sup>12,40</sup> Alcohol use was not a confounder, so was not included in adjusted models. We also did not adjust for blood pressure or the number of prenatal healthcare visits because we considered them potential mediators.

### Statistical analysis

To assess the association of the activity index (quartiles) with birth outcomes, we fit a series of multilevel linear (for birth weight) and logistic (for other outcomes) regression models with random intercepts for mother and community to account for nesting of observations in women and place. The mother-specific intercept incorporated prior pregnancy outcomes (e.g., prior preterm birth) into our models. We selected final models by a combination of *a priori* hypotheses and likelihood ratio tests (P-value < 0.10). For each outcome, model 1 was adjusted for sex of the neonate and season of birth, maternal age at delivery (linear and

quadratic, years), maternal race/ethnicity (white, black, Hispanic, other), primary care status (yes vs. no), smoking status during pregnancy (never, former, current, or conflicting/missing), pre-pregnancy BMI (underweight: z-score > 2SD below mean or < 18.5 kg/m<sup>2</sup>; normal: z-score within 1 SD of mean or 18.5-24.9 kg/m<sup>2</sup>; overweight: z-score 1-2 SD above mean or 25-29.9 kg/m<sup>2</sup>; or obese: z-score > 2 SD above mean or > 30 kg/m<sup>2</sup>), parity (nulliparous vs. multiparous), receipt of Medical Assistance (never vs. ever), delivery hospital (Geisinger Medical Center vs. Geisinger Wyoming Valley), distance to nearest major road in meters, drinking water source (well water vs. municipal), community socioeconomic deprivation (quartiles), and greenness (continuous). In model 2, we further adjusted associations for year (2009-2010 vs. 2011-2013). Birth weight models were also adjusted for gestational age (linear and quadratic, weeks) and high-risk pregnancy models were additionally adjusted for the average annual number of entries on the problem list to account for the fact that its use increased over time (mean of 14% more entries per year).

In sensitivity analyses we included the number of antibiotic orders during pregnancy, restricted preterm models to neonates born moderately to late preterm (32-36 weeks gestation), and fit a Cox proportional hazard model with gestational age as the timescale, preterm birth as the outcome, unconventional natural gas development varying by week, and robust standard errors. We also assessed the possibility of unobserved confounding by assigning babies born in 2006, before there was any unconventional natural gas development, the estimated exposure metric they would have accrued had they been born in 2012, when there was such development. If the 2012 unconventional natural gas development exposure metric were found to be associated with birth outcomes for these 2006 babies, it would suggest that our main study findings may have been spurious.

We report associations as difference in term birth weight or odds ratios for preterm birth, small for gestational age, 5-minute Apgar score, and high-risk pregnancy comparing quartile 2 of unconventional natural gas development activity to quartile 1 with 95% confidence intervals. Models did not exhibit residual spatial variation, which we checked for by visually inspecting semivariograms.<sup>41</sup> Because of the low proportion of missing data (0-1.4% on outcomes and 0-5.2% on confounders) and because missingness only appeared to be associated with year (more missing data in earlier years), patients were omitted from models when they were missing data. We used Stata version 13 (StataCorp. College Station, TX) and R version 3.0.0 (R Foundation for Statistical Computing).

## RESULTS

We identified 20598 neonates born to 20569 mothers who delivered between 2006 and January 2013. After exclusions (Figure 1), we reached a final study sample of 9384 mothers who delivered 10496 neonates (mean of 1.2 per mother). Mothers lived in 699 communities (mean of 14 per community). In eTable 1 we compare the final population to those excluded. Geisinger patients had residential stability. We compared addresses from 2010 and 2013 on 333,222 patients and found that 79.8% had the exact same street address, 6.0% had moved <1500m and another 10% had moved 1500-16,000m from their original address.

The mean birth weight was 3272 grams (SD = 612). Eleven percent (n = 1103) of the births were preterm, 8% were moderately preterm (n = 871), 2% (n = 227) had 5 minute Apgar scores < 7, 10% (n = 1024) were small for gestational age, as expected given our use of an internal standard, and 27% (n = 2853) of pregnancies were identified as high-risk (Table 1).

Unconventional natural gas development in the Pennsylvania Marcellus shale began in the southwest in 2005 (15 wells drilled) and quickly accelerated. By the period 2009-2012, an average of 1555 unconventional wells, drilled to an average depth of 3380m, and 1177 wells entered production annually (Figure 2). The mean (SD), median (IQR) number of wells within 20 km of mothers (during their pregnancy) in the first vs. fourth quartile of exposure to unconventional natural gas development was 6 (28), 0 (0-1) vs. 124 (202), 8 (1-122), respectively, reflecting a marked difference in intensity of potential exposure.

In Table 1 and 2 we present descriptive statistics of several demographic and clinical variables by UNGD activity quartile and by outcome. Neonates born in later years and in the summer and fall; and mothers that were multiparous, received an antibiotic order during pregnancy, used well water, or lived farther from the nearest major road appeared to have higher exposure to unconventional natural gas development activity. Among those with poor pregnancy outcomes, several covariates were more common including receipt of Medical Assistance, black race/ethnicity, ever-smoking, and others (Table 2). Mothers with a primary care provider had an average of 16 prenatal visits (SD = 6) compared to 12 (SD = 7) in those without.

The activity index was not associated with adverse birth outcomes in unadjusted analyses (Table 1). In adjusted birth weight and preterm models, current smoking, underweight BMI, nulliparity, high community socioeconomic deprivation (preterm only), and black race/ethnicity and receipt of Medical Assistance (birth weight only) were positively associated; normal BMI, never smoking, farther distance to nearest major road, and higher residential greenness (preterm only) were negatively associated.

After adjustment for covariates, the fourth quartile of the activity index was associated with lower term birth weight, but not after further adjustment for year (Table 3). In adjusted models, the odds of preterm birth increased across quartiles of the activity index (fourth vs. first quartile, 1.4 [95% CI: 1.0-1.9]) (Table 3). This association strengthened with adjustment for year (Table 3), persisted in a survival model framework (eTable 2), and was robust to restriction to moderate and late preterm births (fourth vs. first quartile, OR = 1.5 [95% CI = 1.0-2.4]). In model 2, antibiotic orders were associated with preterm birth (OR = 1.5 [95% CI = 1.3-1.6]). Unconventional natural gas development exposure during the prenatal period was associated with high-risk pregnancy (fourth vs. first quartile of the activity index, OR = 1.3 [95% CI: 1.1-1.7]), but not with 5 minute Apgar score or small for gestational age (results not shown).

In a sensitivity analysis in infants born in 2006 (n = 1932), future exposure to unconventional natural gas development was not associated with preterm birth, Apgar score, or small for gestational age birth in fully adjusted models. Neonates born in 2006, who

would have been in the 4<sup>th</sup> quartile of the activity index had they been born in 2012, had lower birth weights ( $\beta = -53$  [95% CI -120 to 12]).

## DISCUSSION

We used electronic health record data to conduct a population-based retrospective cohort study in central and northeast Pennsylvania during a time of very rapid unconventional natural gas development in the region. Our study examined associations between prenatal exposure to unconventional natural gas development activity and four birth outcomes and high-risk pregnancy in the mother. We demonstrated that mothers with higher activity index values during pregnancy were more likely to give birth preterm, a finding corroborated in time-to-delivery analysis, and to have a physician-recorded high-risk pregnancy. An association with term birth weight was not robust to adjustment for year. In a sensitivity analysis, when we assigned babies born in 2006 the activity index they would have had if they were born in 2012, unconventional natural gas development was associated with lower birth weight, suggesting that the primary association may have been due, at least in part, to unobserved confounding. There were no associations with Apgar score or small for gestational age. The electronic health record allowed us to carefully ascertain both pregnancy outcomes and confounding variables. We were able to control for other community conditions and exposures, including distance to roadways, source of drinking water, and community socioeconomic deprivation. To our knowledge, this is also the first study to base estimates of unconventional natural gas development activity exposure in relation to health risks on four separate phases of well development.

Three recent reviews summarized evidence linking health and unconventional natural gas development and found it lacking.<sup>1-3</sup> Werner et al. identified only four highly relevant peer-reviewed studies related to these processes and health outcomes: two using self-reported symptoms, one of childhood cancer that may not have adequately accounted for latency, and one of birth outcomes.<sup>21,22,26,30</sup> The only published study dealing with birth outcomes reported that density and proximity of gas wells in Colorado, USA, were associated with two birth defects, but also higher birth weight and lower odds of preterm birth.<sup>26</sup> During the study period, the U.S. Energy Information Administration reported that Colorado produced 28 million cubic meters of natural gas unconventionally and 130 million cubic meters conventionally. We were able to study people living in areas with much higher unconventional natural gas development activity; Pennsylvania produced 58 billion cubic meters of natural gas unconventionally in 2012. A second, unpublished study, compared neonates born to mothers residing within 2.5 km of a spudded well to those living within 2.5 km of a permitted, but not spudded, well.<sup>27</sup> This study reported decreased term birth weight (but did not control for gestational age) and increased small for gestational age and 5 minute Apgar scores < 8, but no association with preterm birth. We too observed associations with Apgar scores < 8, but not < 7, as most prior studies have used, and between unconventional natural gas development and term birth weight when we omitted gestational age.

The unconventional natural gas development process is associated with heterogeneous exposures that last varying amounts of time. We did not have the capability to measure exposures directly. However, we were able to account for the varying durations of the



different phases by using published descriptions and information from our own analysis to assign deliveries activity values in defined windows. This should be an improvement over prior studies, which generally used spud date to identify the start of an exposure assumed to last forever, an incorrect assumption.<sup>26,30</sup> Any bias introduced by errors in the estimation of the durations of development phases is likely to be independent of birth outcomes and thus tend to bias associations towards the null.

There are multiple ways unconventional natural gas development activity could influence birth outcomes. Concerns include impacts on air quality,<sup>1-3</sup> ground and surface water quality,<sup>12</sup> and maternal psychosocial stress from noise, increased traffic volumes, and contextual exposures including social disruption and community livability.<sup>4</sup> For many of these, their associations with birth outcomes have been investigated in other settings.<sup>14,17,37,42</sup> For instance, prior literature suggests that a 10 $\mu\text{g}/\text{m}^3$  increase in exposure to PM<sub>2.5</sub> is associated with a 10% increase in odds of preterm birth and low birth weight.<sup>15,18</sup> There are also several proposed mechanisms linking PM exposure to preterm birth including interference with placental development, inflammation, and increased risk of infection.<sup>18</sup> In our study, mothers with higher activity indices were indeed more likely to receive an antibiotic order during their pregnancy. Neighborhood contextual factors have also been consistently associated with birth outcomes.<sup>43</sup> Women living in communities exposed to unconventional natural gas development likely experience both environmental and social exposures that may have synergistic effects on health.<sup>44</sup> Finally, unmeasured confounding could have contributed to our results; our measure of family SES was binary and did not include education, and we also had no information on occupation.

This study had limitations. In an effort to assign activity values more accurately than prior studies, we estimated the duration of each phase of unconventional natural gas development. This is likely to have introduced measurement error since the amount of time each phase lasts varies by well. We used a distance-based metric to estimate exposure to four phases of development, but were not able to evaluate phase-specific associations due to collinearity. Phases are known to contribute different types of exposures (e.g., pad development is a source of diesel emissions including PM as well as noise),<sup>1</sup> but our methodology did not allow us to differentiate among phase-specific exposures, type of hazardous exposure (e.g., air and water pollution), and the contextual effects of development. We were not able to take environmental samples, which may have led to exposure misclassification and prevented us from determining if a specific pollutant was responsible for our associations. Additionally, unconventional natural gas development was highly correlated with year, making it challenging to control for temporal trends; therefore we presented results both unadjusted and adjusted for year. In regards to conventional gas development in the state, although the densest development is in the northwest and many of these wells are decades old and non-producing, there was still collinearity between our activity index and conventional gas proximity metrics, which precluded adjustment for conventional gas well locations. Historical addresses are not retained in the Geisinger electronic health record so we were not able to determine whether the last recorded address represented residential location during the course of pregnancy. Our sensitivity analysis suggested that most Geisinger patients do not move, and if they do, they tend to move locally. In our study, many wells were developed in one location over time, so the exposures, emissions, and community

circumstances present in one trimester were likely present in another. This collinearity prevented us from evaluating trimester-specific associations.

Prior studies found elevated symptoms in regions with unconventional natural gas development and concern by residents of possible health effects. This study adds to limited evidence that unconventional natural gas development adversely affects birth outcomes. We observed that an index of development activity was associated with both preterm birth and high-risk pregnancy. Multiple aspects of development might be involved, including hazardous exposures and contextual effects. Future studies should use direct environmental sampling to more accurately capture exposure and include data on mother's place of residence throughout pregnancy. Such data is needed to allow policy makers to effectively weigh the risks and benefits of unconventional natural gas development.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

We thank Joseph J. DeWalle, Jennifer K. Irving, Joshua M. Crisp, and Awathif Mackeen at the Geisinger Center for Health Research (Danville, PA, USA) for patient geocoding, assistance in assembling the UNGD dataset, and valuable comments on the manuscript; and Peter James at the Harvard School of Public Health (Boston, Massachusetts) and Chih-Da (John) Wu at the National Chiayi University (Chiayi City, Taiwan) for providing the NDVI data.

### Source of Funding:

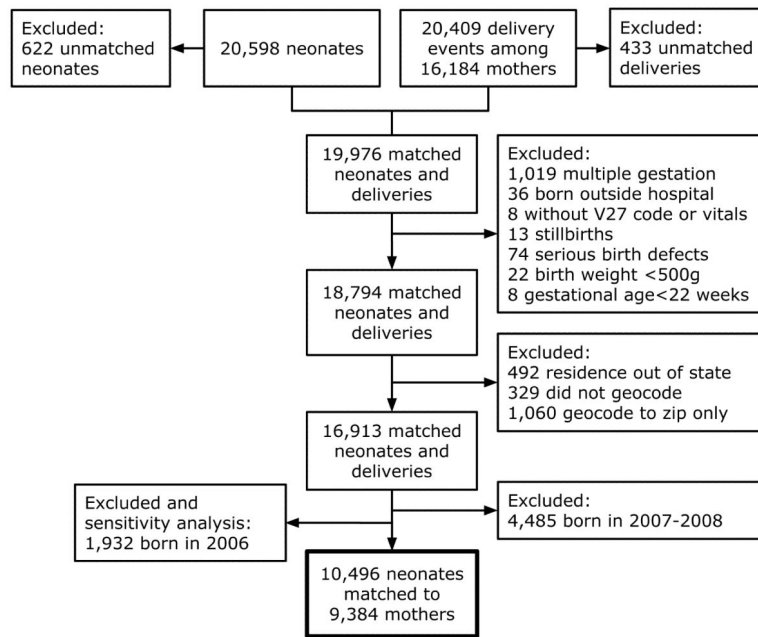
This study was funded by the National Institute of Environmental Health Sciences grant ES023675-01 (PI: B Schwartz) and training grant ES07141 (S Rasmussen). Additional support was provided by the Degenstein Foundation for compilation of well data, the Robert Wood Johnson Foundation Health & Society Scholars program (J Casey), and the National Science Foundation Integrative Graduate Education and Research Traineeship (S Rasmussen).

## REFERENCES

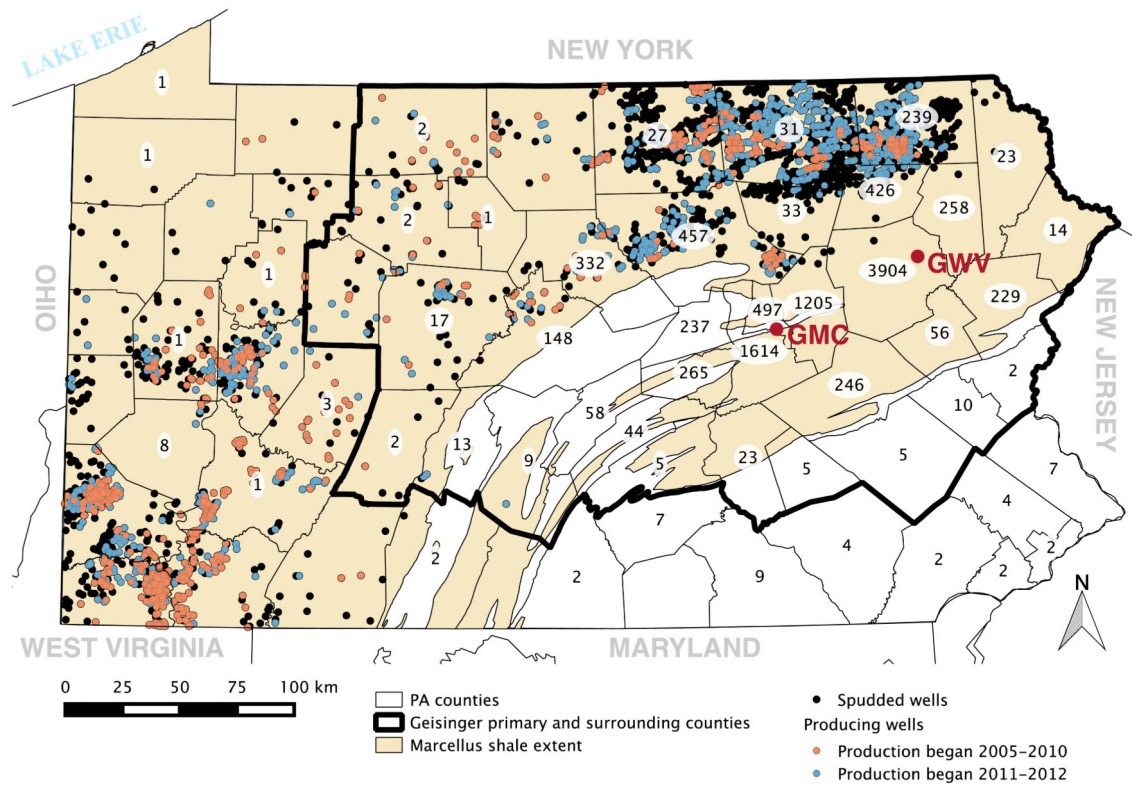
1. Adgate JL, Goldstein BD, McKenzie LM. Potential public health hazards, exposures and health effects from unconventional natural gas development. *Environ Sci Technol*. 2014; 48(15):8307–20. [PubMed: 24564405]
2. Shonkoff SB, Hays J, Finkel ML. Environmental public health dimensions of shale and tight gas development. *Environ Health Perspect*. 2014; 122(8):787–95. [PubMed: 24736097]
3. Werner AK, Vink S, Watt K, Jagals P. Environmental health impacts of unconventional natural gas development: A review of the current strength of evidence. *Sci Total Environ*. 2015; 505C:1127–41. [PubMed: 25461113]
4. Witter RZ, McKenzie L, Stinson KE, Scott K, Newman LS, Adgate J. The use of health impact assessment for a community undergoing natural gas development. *Am J Public Health*. 2013; 103(6):1002–10. [PubMed: 23597363]
5. McKenzie LM, Witter RZ, Newman LS, Adgate JL. Human health risk assessment of air emissions from development of unconventional natural gas resources. *Sci Total Environ*. 2012; 424:79–87. [PubMed: 22444058]
6. Macey GP, Breech R, Chernaik M, et al. Air concentrations of volatile compounds near oil and gas production: a community-based exploratory study. *Environ Health*. 2014; 13:82. [PubMed: 25355625]

7. Roy AA, Adams PJ, Robinson AL. Air pollutant emissions from the development, production, and processing of Marcellus Shale natural gas. *J Air Waste Manag Assoc.* 2014; 64(1):19–37. [PubMed: 24620400]
8. Vinciguerra T, Yao S, Dadzie J, et al. Regional air quality impacts of hydraulic fracturing and shale natural gas activity: Evidence from ambient VOC observations. *Atmos Environ.* 2015; 110(0):144–50.
9. Paulik LB, Donald CE, Smith BW, et al. Impact of Natural Gas Extraction on PAH Levels in Ambient Air. *Environ Sci Technol.* 2015; 49(8):5203–10. [PubMed: 25810398]
10. Kassotis CD, Tillitt DE, Davis JW, Hormann AM, Nagel SC. Estrogen and androgen receptor activities of hydraulic fracturing chemicals and surface and ground water in a drilling-dense region. *Endocrinology.* 2014; 155(3):897–907. [PubMed: 24424034]
11. Jackson RB, Vengosh A, Darrah TH, et al. Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction. *Proc Natl Acad Sci U S A.* 2013; 110(28):11250–5. [PubMed: 23798404]
12. Llewellyn GT, Dorman F, Westland JL, et al. Evaluating a groundwater supply contamination incident attributed to Marcellus Shale gas development. *Proc Natl Acad Sci U S A.* 2015
13. Graham J, Irving J, Tang X, et al. Increased traffic accident rates associated with shale gas drilling in Pennsylvania. *Accident Anal Prev.* 2015; 74(0):203–9.
14. Stieb DM, Chen L, Eshoul M, Judek S. Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. *Environ Res.* 2012; 117:100–11. [PubMed: 22726801]
15. Davvand P, Parker J, Bell ML, et al. Maternal exposure to particulate air pollution and term birth weight: a multi-country evaluation of effect and heterogeneity. *Environ Health Perspect.* 2013; 121(3):267–373. [PubMed: 23384584]
16. Pedersen M, Giorgis-Allemand L, Bernard C, et al. Ambient air pollution and low birthweight: a European cohort study (ESCAPE). *Lancet Respir Med.* 2013; 1(9):695–704. [PubMed: 24429273]
17. Davvand P, Basagana X, Figueras F, et al. Air pollution and preterm premature rupture of membranes: a spatiotemporal analysis. *Am J Epidemiol.* 2014; 179(2):200–7. [PubMed: 24125920]
18. Pereira G, Belanger K, Ebisu K, Bell ML. Fine particulate matter and risk of preterm birth in Connecticut in 2000–2006: a longitudinal study. *Am J Epidemiol.* 2014; 179(1):67–74. [PubMed: 24068199]
19. Saberi P, Probert KJ, Powers M, Emmett E, Green-McKenzie J. Field survey of health perception and complaints of Pennsylvania residents in the Marcellus Shale region. *Int J Environ Res Public Health.* 2014; 11(6):6517–27. [PubMed: 25003172]
20. Rabinowitz PM, Slizovskiy IB, Lamers V, et al. Proximity to natural gas wells and reported health status: results of a household survey in Washington County, Pennsylvania. *Environ Health Perspect.* 2015; 123(1):21–6. [PubMed: 25204871]
21. Perry SL. Using ethnography to monitor the community health implications of onshore unconventional oil and gas developments: examples from Pennsylvania's Marcellus Shale. *New Solut.* 2013; 23(1):33–53. [PubMed: 23552647]
22. Steinzor N, Subra W, Sumi L. Investigating links between shale gas development and health impacts through a community survey project in Pennsylvania. *New Solut.* 2013; 23(1):55–83. [PubMed: 23552648]
23. Casey JA, Ogburn EL, Rasmussen SG, et al. Predictors of Indoor Radon Concentrations in Pennsylvania, 1989–2013. *Environ Health Perspect.* 2015
24. Davvand P, Ostro B, Figueras F, et al. Residential proximity to major roads and term low birth weight: the roles of air pollution, heat, noise, and road-adjacent trees. *Epidemiology.* 2014; 25(4): 518–25. [PubMed: 24787556]
25. Staneva A, Bogossian F, Pritchard M, Wittkowski A. The effects of maternal depression, anxiety, and perceived stress during pregnancy on preterm birth: A systematic review. *Women and Birth.* 2015
26. 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 Perspect.* 2014; 122(4):412–7. [PubMed: 24474681]

27. Hill, EL. Shale Gas Development and Infant Health: Evidence from Pennsylvania. 2013. <http://dyson.cornell.edu/research/researchpdf/wp/2012/Cornell-Dyson-wp1212.pdf> (accessed 5/15/2015)
28. Kovats S, Depledge M, Haines A, et al. The health implications of fracking. *Lancet*. 2014; 383(9919):757–8. [PubMed: 24581655]
29. Mitka M. Rigorous evidence slim for determining health risks from natural gas fracking. *JAMA*. 2012; 307(20):2135–6. [PubMed: 22618904]
30. Fryzek J, Pastula S, Jiang X, Garabrant DH. Childhood cancer incidence in Pennsylvania counties in relation to living in counties with hydraulic fracturing sites. *J Occup Environ Med*. 2013; 55(7): 796–801. [PubMed: 23836020]
31. Schwartz BS, Stewart WF, Godby S, et al. Body mass index and the built and social environments in children and adolescents using electronic health records. *Am J Prev Med*. 2011; 41(4):e17–e28. [PubMed: 21961475]
32. Casey, JA. Spatial and molecular characterization of methicillin-resistant *Staphylococcus aureus* (MRSA) infection in relation to high-density livestock production in Pennsylvania. Johns Hopkins Bloomberg School of Public Health; Baltimore, MD: 2014. PhD Dissertation
33. Stuart A, Olausson PO, Källén K. Apgar scores at 5 minutes after birth in relation to school performance at 16 years of age. *Obstet Gynecol*. 2011; 118(2):201–8. Part 1. [PubMed: 21734618]
34. New York State Department of Environmental Conservation. Final Supplemental Generic Environmental Impact Statement on Oil, Gas, and Solution Mining Regulatory Program. 2015. [http://www.dec.ny.gov/docs/materials\\_minerals\\_pdf/fsgeis2015.pdf](http://www.dec.ny.gov/docs/materials_minerals_pdf/fsgeis2015.pdf) (accessed 5/15/15)
35. Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. *Lancet*. 2008; 371(9606):75–84. [PubMed: 18177778]
36. Valero De Bernabe J, Soriano T, Albaladejo R, et al. Risk factors for low birth weight: a review. *Eur J Obstet Gynecol Reprod Biol*. 2004; 116(1):3–15. [PubMed: 15294360]
37. Padula AM, Mortimer K, Hubbard A, Lurmann F, Jerrett M, Tager IB. Exposure to traffic-related air pollution during pregnancy and term low birth weight: estimation of causal associations in a semiparametric model. *Am J Epidemiol*. 2012; 176(9):815–24. [PubMed: 23045474]
38. Vos AA, Posthumus AG, Bonsel GJ, Steegers EA, Denktas S. Deprived neighborhoods and adverse perinatal outcome: a systematic review and meta-analysis. *Acta Obstet Gynecol Scand*. 2014; 93(8):727–40. [PubMed: 24834960]
39. Hystad P, Davies HW, Frank L, et al. Residential greenness and birth outcomes: evaluating the influence of spatially correlated built-environment factors. *Environ Health Perspect*. 2014; 122(10):1095–102. [PubMed: 25014041]
40. Manassaram DM, Backer LC, Moll DM. A review of nitrates in drinking water: maternal exposure and adverse reproductive and developmental outcomes. *Ciencia & saude coletiva*. 2007; 12(1): 153–63. [PubMed: 17680066]
41. Cressie, N. *Statistics for Spatial Data*. J. Wiley & Sons, Inc; 1991.
42. Gehring U, Tamburic L, Sbihi H, Davies HW, Brauer M. Impact of noise and air pollution on pregnancy outcomes. *Epidemiology*. 2014; 25(3):351–8. [PubMed: 24595395]
43. Metcalfe A, Lail P, Ghali WA, Sauve RS. The association between neighbourhoods and adverse birth outcomes: a systematic review and meta-analysis of multi-level studies. *Paediatr Perinat Epidemiol*. 2011; 25(3):236–45. [PubMed: 21470263]
44. Erickson AC, Arbour L. The shared pathoetiological effects of particulate air pollution and the social environment on fetal-placental development. *J Environ Public Health*. 2014; 2014:901017. [PubMed: 25574176]



**Figure 1.**  
Flow diagram of study population assembly



**Figure 2.**

The Marcellus shale extent, the location of spudded and producing wells as of December 2012, the location of the two Geisinger Health System hospitals and the primary and surrounding Geisinger counties. Annotation indicates the number of neonates born to mothers residing in each county. GMC = Geisinger Medical Center. GWV = Geisinger Wyoming Valley.

**Table 1**

Distribution of study population characteristics among 9384 mothers and their 10496 children by quartile of unconventional natural gas development (UNGD) activity index

Variable	No. (%)	UNGD activity index quartile <sup>a</sup>			
		1	2	3	4
<b>Maternal characteristics</b>					
Age at birth, years, mean (SD)	10496 (100)	27.6 (5.8)	27.8 (5.7)	27.9 (5.7)	27.8 (5.8)
Race/ethnicity, %					
White	9327 (89)	88	89	86	92
Black	382 (4)	4	3	4	3
Hispanic	601 (6)	6	6	7	3
Other	148 (1)	2	1	2	1
Missing	38 (<1)	<1	<1	<1	<1
Primary care patient, %	4789 (46)	45	45	46	46
Smoking status <sup>b</sup> , %					
Never	4984 (47)	46	45	49	49
Former	2258 (22)	21	24	21	20
Current	1785 (17)	18	18	15	17
Conflicting or missing	1489 (14)	15	13	15	14
Alcohol use during pregnancy <sup>b</sup> , %					
No	8448 (80)	77	79	83	83
Yes	1412 (13)	14	14	13	13
Missing	636 (6)	9	7	4	4
Pre-pregnancy body-mass index (kg/m <sup>2</sup> ), %					
<18.5	222 (2)	2	2	2	2
18.5-24.9	3878 (37)	37	38	36	36
25-29.9	2834 (27)	27	25	28	28
30	3013 (29)	29	30	28	28
Missing	549 (5)	5	5	5	5
Pre-pregnancy blood pressure, %					
Systolic >140mmHg or diastolic >90mmHg	1125 (11)	9	11	13	10
Normal	9371 (89)	91	89	87	90
Nulliparous, %	4600 (44)	47	43	44	41
Healthcare visits during pregnancy, n, mean (SD)	10496 (100)	14.4 (6.3)	13.8 (6.4)	13.6 (6.7)	13.7 (6.7)
Antibiotic order during pregnancy, %	3338 (32)	30	31	31	35
Receipt of Medical Assistance, %	4796 (46)	44	47	45	47
Delivery hospital, %					
Geisinger Medical Center	5638 (54)	57	57	51	49

Variable	No. (%)	UNGD activity index quartile <sup>a</sup>			
		1	2	3	4
Geisinger Wyoming Valley	4858 (46)	43	43	49	51
Distance to nearest major road, m, median (IQR)	10496 (100)	788 (284-2825)	863 (304-3229)	609 (237-1826)	1373 (455-6757)
Drinking water source, %					
Municipal water	7306 (70)	72	72	78	57
Well water	3190 (30)	28	28	22	43
Community socioeconomic deprivation <sup>c</sup> , %					
Quartile 1	2590 (25)	25	23	24	27
Quartile 2	2648 (25)	23	22	23	28
Quartile 3	2642 (25)	25	23	24	29
Quartile 4	2616 (25)	27	33	29	15
Residential greenness, NDVI index, mean (SD)	0.54 (0.10)	0.50 (0.11)	0.56 (0.09)	0.54 (0.09)	0.54 (0.11)
<b>Infant Characteristics</b>					
Male, %	5372 (51)	51	52	52	50
Birth weight, grams, mean (SD)	10495 (100)	3289 (604)	3249 (623)	3286 (599)	3264 (622)
Gestational age, weeks, mean (SD)	10418 (99)	38.9 (2.2)	38.9 (2.4)	39.0 (2.1)	38.9 (2.3)
Preterm birth <37 weeks, %	1103 (11)	10	11	10	11
Preterm birth 32 to 36 weeks, %	871 (8)	2	2	2	2
Small for gestational age, %	1024 (10)	9	10	10	10
Apgar score, %					
5 minute, <7	227 (2)	2	2	2	2
5 minute, 7	10199 (95)	97	97	97	97
5 minute, missing	70 (<1)	1	<1	1	1
High-risk pregnancy <sup>d</sup> , %	2853 (27)	17	25	33	33
Birth year, %					
2009	2336 (22)	79	7	1	2
2010	2518 (24)	20	55	9	11
2011	2608 (25)	1	27	49	22
2012	2852 (27)	<1	11	38	60
2013	182 (2)	0	<1	2	5
Birth season, %					
December-February	2562 (24)	27	20	25	24
March-May	2605 (25)	29	25	24	21
June-August	2748 (26)	23	29	25	27
September-November	2581 (25)	20	26	25	27

UNGD activity index quartile was assigned based on 4 z-transformed indicators using inverse-distance squared models that incorporated distance to the mother's home; dates and durations of the phases (well pad development, spudding, hydraulic fracturing, and production); and well characteristics (depth and production volume) during gestation, and is in standard deviation units. Percentages are rounded to whole numbers.

EHR = electronic health record. IQR = interquartile range. NDVI = normalized difference vegetation index.



<sup>a</sup> Quartile 1: <-0.44; Quartile 2: -0.43 to -0.15, Quartile 3: -0.14 to 0.18, Quartile 4: >0.18.

<sup>b</sup> Smoking, alcohol use, and high-risk pregnancy were reported during pregnancy in the EHR social history and problem list.

<sup>c</sup> Community socioeconomic deprivation was assigned at the township, borough, or census tract level, based on 6 indicators derived from the U.S. Census American Community Survey 2012 5-year estimates: combined less than high school education, not in the labor force, in poverty, on public assistance, civilian unemployment, and does not own a car; a higher score represents a more deprived community.

<sup>d</sup> Defined based on physician-reported high-risk pregnancy.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2**

Distribution of outcomes by selected covariates

	Outcome				
	Birth weight, g, median (IQR)	Preterm birth, n (%)	5 min Apgar <7, n (%)	SGA, n (%)	High risk pregnancy <sup>a</sup> , n (%)
<b>N</b>	10495	1103	10426	1024	2853
<b>Pre-pregnancy body-mass index (kg/m<sup>3</sup>)</b>					
<18.5	3051 (2696-3359)	50 (23)	7 (3)	41 (19)	66 (30)
18.5-24.9	3258 (2903-3575)	408 (11)	80 (2)	443 (12)	1008 (26)
25-29.9	3352 (2991-3685)	265 (9)	66 (2)	267 (10)	751 (26)
30	3404 (3071-3745)	286 (10)	57 (2)	222 (7)	940 (31)
Missing	3263 (2908-3631)	94 (17)	17 (3)	51 (10)	89 (16)
<b>Parity</b>					
Nulliparous	3303 (2940-3625)	486 (11)	116 (2)	525 (12)	981 (21)
Multiparous	3338 (2991-3686)	617 (10)	111 (2)	499 (9)	1872 (32)
<b>Antibiotic order during pregnancy</b>					
No	3348 (3012-3679)	580 (8)	131 (2)	686 (10)	1891 (26)
Yes	3268 (2885-3617)	523 (16)	96 (3)	338 (10)	962 (29)
<b>Year of birth</b>					
2009 and 2010	3330 (2974-3665)	528 (11)	90 (2)	455 (10)	888 (18)
2011, 2012, and 2013	3314 (2968-3657)	575 (10)	138 (2)	569 (10)	1965 (35)
<b>Delivery hospital</b>					
Geisinger Medical Center	3284 (2884-3630)	874 (16)	180 (3)	554 (10)	1507 (27)
Geisinger Wyoming Valley	3365 (3050-3688)	229 (5)	47 (1)	470 (10)	1346 (28)
<b>Community socioeconomic deprivation<sup>b</sup></b>					
Quartile 1	3372 (3033-3700)	249 (10)	67 (3)	205 (8)	597 (23)
Quartile 2	3345 (2984-3667)	264 (10)	49 (2)	241 (9)	705 (27)
Quartile 3	3303 (2944-3640)	306 (12)	53 (2)	262 (10)	727 (28)
Quartile 4	3264 (2925-3620)	284 (11)	58 (2)	316 (12)	824 (32)

Percentages are rounded to whole numbers.

EHR = electronic health record. IQR = interquartile range. SGA = small for gestational age.

<sup>a</sup>Reported in EHR problem list during pregnancy.<sup>b</sup>Community socioeconomic deprivation was assigned at the township, borough, or census tract level, based on 6 indicators derived from the US Census American Community Survey 2012 5-year estimates: combined less than high school education, not in the labor force, in poverty, on public assistance, civilian unemployment, and does not own a car; a higher score represents a more deprived community.

**Table 3**

Associations of term birth weight and preterm birth and exposure to unconventional natural gas development (UNGD) activity

	Model 1A <sup>a</sup>	Model 2A <sup>b</sup>	Model 1B <sup>c</sup>	Model 2B <sup>d</sup>
	Term birth weight (g)		Preterm birth	
Variable	Difference (95% CI)	Difference (95% CI)	OR (95% CI)	OR (95% CI)
UNGD activity quartile	N = 8839	N = 8839	N = 9848	N = 9848
1	Reference	Reference	1.0	1.0
2	-21 (-46 to 5)	-16 (-44 to 11)	1.2 (0.9-1.6)	1.3 (1.0-1.8)
3	-9 (-35 to 16)	1 (-34 to 36)	1.3 (1.0-1.7)	1.6 (1.1-2.4)
4	-31 (-57 to -5)	-20 (-56 to 16)	1.4 (1.0-1.9)	1.9 (1.2-2.9)
Year of birth				
2009 or 2010		Reference		1.0
2011, 2012, or 2013		12 (-15 to 39)		1.3 (1.0-1.8)

CI=confidence interval. OR = odds ratio.

<sup>a</sup> Model 1A was adjusted for sex and gestational age of neonate; maternal characteristics: age at delivery, race/ethnicity, primary care patient status, smoking status, pre-pregnancy body mass index, parity, number of antibiotic orders during pregnancy, receipt of Medical Assistance, delivery hospital, drinking water source, distance to nearest major road, mean residential greenness during pregnancy; and community socioeconomic deprivation quartile.

<sup>b</sup> Model 2A further adjusted for year of birth.

<sup>c</sup> Model 1B was adjusted for sex of neonate; maternal characteristics: age at delivery, race/ethnicity, primary care patient status, smoking status, pre-pregnancy body mass index, parity, receipt of Medical Assistance, delivery hospital, drinking water source, distance to nearest major road, mean residential greenness during pregnancy; and community socioeconomic deprivation quartile.

<sup>d</sup> Model 2B further adjusted for year of birth.