Assessing exposure to unconventional natural gas development: Using an air pollution dispersal screening model to predict new-onset respiratory symptoms

David R. Brown[1], Lydia H. Greiner[1,2*], Beth I. Weinberger[1], Leslie Walleigh[1], and Dale Glaser[2]

[1]*Southwest Pennsylvania Environmental Health Project, McMurray, PA, USA*

[2]*San Diego State University, San Diego, CA, USA*

Abstract

Various exposure estimates have been used to assess health impact of unconventional natural gas development (UNGD). The purpose of this study was to (1) use an air pollution dispersal screening model and wind direction to characterize the air emissions from UNGD facilities at each residence and (2) assess association of this exposure estimate with respiratory symptoms. Respiratory symptoms were abstracted from health records of a convenience sample of 104 adults from one county in southwestern PA who had completed a standard clinical interview with a nurse practitioner. Using publicly available air emission data, we applied a “box” air pollution dispersion screening model to estimate the median ambient air level of CO, NO_x, PM 2.5, VOCs, and formaldehyde at the residence during the year health symptoms were reported. Sources and median emissions were categorized as north, south, east, or west of the residence to account for the effect of wind direction on dispersion. Binary logistic regression was performed for each respiratory symptom. Number of sources had varying magnitudes of association with some symptoms (i.e., cough, shortness of breath, and “any respiratory symptom”) and no association with others (i.e., sore throat, sinus problems, wheezing). Air emissions were not associated with any symptom.

---

*Address correspondence to Lydia Greiner, Southwest Pennsylvania Environmental Health Project, 2001 Waterdam Plaza Drive, Suite 201, McMurray, PA 15317; Phone: 203-767-1598; E-mail: lgreiner@environmentalhealthproject.org

KEYWORDS: Hydraulic fracking, unconventional natural gas, air emissions, exposure, health.

Introduction

Over the past decade many areas of rural Pennsylvania have become industrialized by the rapid expansion of unconventional natural gas development (UNGD). This development has resulted in thousands of point sources of hazardous air emissions distributed over the previously rural landscape. In spite of the well documented harmful health effects of some of the most abundant emissions from UNGD, such as particulate matter, nitrogen oxides, and volatile organic compounds, [1, 2] the overall public health impact of UNGD has been difficult to assess due to limited exposure information.

Epidemiologic studies have looked at the associations between UNGD exposure and birth outcomes; [3-5] childhood cancer; [6] a constellation of symptoms commonly reported: headache, fatigue, and rhinosinusitis; [7] vehicular accidents; [8] depressive symptoms; [9] and asthma exacerbations. [10] Self-report studies have documented an association between respiratory symptoms and proximity to UNGD activity. [11]

Our previous work described the prevalence of self-reported symptoms in a sample of adults who lived within two kilometers of at least one unconventional natural gas well in southwestern PA. [12] Five of the 10 most frequently reported symptoms were respiratory: sore throat (39%); cough (33%); shortness of breath (29%); sinus problems (29%); and wheezing (22%). This self-reported prevalence of respiratory symptoms in residents of communities exposed to UNGD is consistent with the work of Rabinowitz et al., [11] Rasmussen et al., [10] and Tustin et al. [7] Rabinowitz et al. found a relationship between respiratory symptoms and increased proximity to UNGD wells; Rasmussen et al. and Tustin et al. found relationships between exposure and asthma exacerbations and sinusitis, respectively, using an exposure metric incorporating four stages of well development (pad preparation, drilling, stimulation, and production) along with the sum of the inverse squared distances between wells and residences.

One of the challenges of looking at health outcomes related to exposure is that it is unclear how best to characterize residential exposure to UNGD emissions. Different metrics, virtually all of which have focused on wells while ignoring potential contribution of other infrastructure, have been used to estimate exposure. Among the simplest is distance to unconventional gas wells. [11, 13] Others have used the inverse distance weighted well count to account for wells over a large distance, giving greater weight to those closest to the residence. [3, 4] Responding to the variation in emissions at different stages of well development, still others have incorporated those stages into metrics that included distance and number and depth of wells and production of gas. [5, 7, 9, 10] Koehler et al. [14] note wells are just a part of the infrastructure of UNGD. Additional sources of exposure include impoundments, where hydraulic fracturing fluid that returns to the surface from the well is stored; flaring events, where excess gas is burned off; compressor stations, that keep the gas moving through the pipelines; and pipelines themselves, that carry the gas from processing plants to the end user. Underscoring the importance of including as many recognized sources of air emissions as feasible, they found that a metric that incorporated both proximity to four stages of well development and, additionally, to compressor stations was more predictive of mild asthma exacerbations than either of the two other metrics used in prior epidemiological studies (i.e., inverse weighted distance or proximity to four stages of well development). Although Koehler et al. initially intended to incorporate data on flaring events and impoundments, they found that the data on these two important sources of air emissions were too sparse.

We have developed two measures of exposure for this analysis to better capture sources of emissions and the impact of air movement on exposure. The goals of this analysis were to (1) characterize the UNGD-related emissions from well pads, processing plants, and compressor stations; (2) use an air pollution dispersal screening model and wind direction to estimate household-level exposure to these emissions; and (3) test the relationship between the exposure measures and respiratory health outcomes.

Materials and methods

Health outcomes

As previously described in detail, [12] the Southwest Pennsylvania Environmental Health Project has been systematically collecting health data from residents of communities located near UNGD sites since 2012. Between February 1, 2012 and December 31, 2017, 164 adults and children completed the standardized health assessment, typically conducted face-to-face by a family nurse practitioner using standard clinical practice for collecting a medical history, current problems, review of systems, past medical history, family history, and social history. These individuals self-selected to complete the health assessment with the nurse practitioner because of concerns about symptoms they were experiencing. The 164 records in this convenience sample were reviewed retrospectively by a team of health care providers that included a physician who is board certified occupational medicine (LW) and at least one nurse practitioner.

For this analysis, exclusion criteria included: age less than 18; employment in the oil or gas industry; incomplete health assessment; and residence outside of the county of interest. After exclusion criteria were applied, a convenience sample of 104 records was available for this analysis. Symptoms were abstracted from each record; symptoms were excluded if they could plausibly be attributed to pre-existing or current health conditions (e.g., chronic obstructive pulmonary disease in the case of “shortness of breath”) or behaviors (e.g., tobacco smoking in the case of “cough”). Five respiratory symptoms were abstracted and used as dichotomous outcome in this analysis: cough, shortness of breath, sinus problems, sore throat, and wheezing. A sixth dichotomous outcome, “any respiratory symptom”, was used to indicate the presence of at least one respiratory symptom.

Exposure measures

Since March 31, 2012, owners and operators of natural gas production and processing operations have been required to report air emissions to the Pennsylvania Department of Environmental Protection (PA DEP) annually. Initially, owners and operators were required to report emissions from the following sources: stationary engines, heaters, tanks/impoundments, dehydration units, pneumatic pumps, fugitives, venting and blowdown, drill rigs, and well completions; in 2013, compressor stations were added to the list of sources. Air contaminants reported from these sources include: carbon monoxide, nitrogen oxides, particulate matter less than 10 micrometers in diameter (PM10), particulate matter less than 2.5 micrometers in diameter (PM2.5), sulfur dioxide, volatile organic compounds, and additional hazardous air pollutants including benzene, ethylbenzene, formaldehyde, N-hexane, toluene and 2,2,4-trimethylpentane.

Greenhouse gas emissions are also reported. [15] The publicly available annual Air Emission Inventory reports “tons per year” for each compound released from each source. [16] It is important to note that, while emissions from each type of source have some shared characteristics, they also differ in kind and quantity. [17]

Data from 2012-2016 were available. [16] We used the annual Air Emissions Inventory that corresponded to the year the health assessment was collected. Data from the Air Emissions Inventory that corresponded to the year of the health assessment were not available for 17 of the 104, reducing our sample to 87. Geocoding was used to determine the proximity of the sources listed in the Air Emissions Inventory to each residence. Latitude and longitude coordinates for each source are published in the Inventory. Addresses, available for each residence, were converted into latitude and longitude. [18] Using these coordinates, we identified all sources reported on the annual Air Emissions Inventory that were located within two kilometers of each residence. For our exposure estimate, we selected the five compounds with the highest reported mass and known health effects: nitrogen oxides, carbon monoxide, VOCs, PM 2.5, and formaldehyde. While the greenhouse gasses methane and carbon dioxide are among those compounds with the highest reported mass they were not included as they do not have acute health effects at levels typical for environmental exposures. Emissions data were reported for each compound as tons per year. For this analysis we converted tons/year into grams per hour.

There are limitations to the PA DEP data. Not all sources of emissions are included in the report. Emissions from the largest Title V compressor stations and selected compressor stations along interstate pipelines are reported only to the federal government and as a result are not included in the Air Emissions Inventory. Air emissions resulting from flaring events or evaporation from impoundments are not included in the Air Emissions Inventory, and indoor exposures to harmful compounds off gassed from contaminated water remain the subject of conjecture. Air emissions are reported as “tons per year”; our conversion to “grams per hour” assumes a consistent rate of emissions and dilutes the potential effect of peak exposures.

Patterns of dispersal of air emissions are influenced by weather and atmospheric conditions. We estimated exposure at a residence using a “box” air pollution dispersion screening model, based on the work of Pasquill. [19, 20] Our approach, described in detail elsewhere, [17, 21] will be described briefly here. The “box” air pollution dispersion screening model is based on a theoretical box (volume of air) that carries emissions from a source. That box increases in size (dispersion) based on parameters that determine air dilution down-wind from emission sources to estimate the ambient air level of pollutants. The parameters are: 1) cloud cover which determines vertical mixing due to heating and cooling of ground surface; 2) wind speed which determines time for horizontal mixing; and 3) time of day, which influences stability of the air system. Using these parameters, Pasquill developed a model with six stability classes defined by particular combinations of atmospheric conditions, as shown in Table 1.

The initial volume of the box is calculated by

$$Volume\ (m^2)=a*b*c$$

where a = meters of air that pass over a site/minute; b = 100 meters (assumed dimension of typical site); and c = intercept of stability class and distance from source (see Figure 1).

The air concentration of a compound at the source is calculated by

$$S\ (ug/m^3) = Emissions\ /\ Volume$$

where *Emissions*= mass in ug/minute and *Volume=a*b*c*.

To determine the dilution downwind from the source the volume of the box at specific distances from the source must be recalculated by

$$VOLUME\ (m^2)=a*B*C$$

where a = meters of air that pass over a site/minute; B = expansion of the box (see Figure 1); and C = intercept of stability class and distance from source (see Figure 1).

We applied this model using data from the National Oceanic and Atmospheric Association (NOAA). [22] NOAA provides hourly cloud cover, wind speed, time of day, and wind direction. For this analysis, we used data collected at the Allegheny County Airport in West Mifflin, PA. This station is located on average 20-30 miles from the residences included in this analysis. Data from the West Mifflin station were used to establish hourly air stability classes over the year at each residence in our sample.

For each source within two km of a residence, we used the screening model to calculate the hourly ambient air levels at the source and at the residence in ug/m^3 for CO, NO_x, PM 2.5, VOCs, and formaldehyde emitted over the year. We determined the mean, standard deviation, and median of the hourly air emissions for each compound and then summed the medians. Table 2 shows the results of this screening model applied to a hypothetical source emitting 300 grams/hour. As shown on Table 2, we used the model to calculate ambient air levels at seven unique distances ranging from 0.1 km (i.e., “fence line”) to 10 km from the hypothetical source, with air emissions expressed in ug/m^3. [20]

The “box” air pollution dispersion screening model does not take in to account wind direction. Our examination of data from the National Oceanic and Atmospheric Association [22] revealed that wind direction is from the north, south, and west approximately 90% of the time (i.e., 30% of the time from each of these directions), and from the east approximately 10% of the time. To account for the variation in wind, each source was categorized as north, south, east, or west of the residence, based on distance from “true”. For example, we considered a source north of a residence if it was within 45 degrees of true north, and the same for the other three directions.

As an estimate of exposure to air emissions we generated eight variables for this analysis. For each residence we generated four variables (“north sources”, “south sources”, “east sources”, and “west sources”) which represented the number of sources in each quadrant within two km of the residence. We generated four additional variables (“north emissions”, “south emissions”, “east emissions”, and “west emissions”) which represented the median of ambient air levels of emissions from wells, processing plants, and compressor stations in the quadrant.

Analytic plan

Given the nature of the dichotomous outcomes (i.e., any respiratory symptom, cough, shortness of breath, sore throat, sinus problems, and wheezing), binary logistic regression was performed. [23] The quantitative predictors were four emissions (i.e., north, south, east, and west EM) and four sources (i.e., north, south, east, and west source). Moreover, gender, age, and household water source were included in the full model. All parameter estimates (e.g., the logit *b*, odds ratio (OR) and 95% confidence interval, etc.) are reported ($\alpha = .05$), as well as classification statistics such as sensitivity (true positive hit rate) and specificity (true negative hit rate). As well, pseudo r^2 statistics are reported, including the Cox-Snell and Nagelkerke R^2, each of which is a function of the -2LL statistic for the full and restricted model (LL = log likelihood). The Hosmer-Lemeshow test was used to assess model fit [24] and for this test statistic, non-significance is preferred (i.e., expected probabilities approximate observed probabilities).

Results and discussion

The median age of the 87 adults in this convenience sample was 57 (SD 12); 60% were female; and 40% reported using well water for human activities such as cooking, drinking, and/or bathing.

UNGD-related emissions from well pads, processing plants, and compressor stations

There are 16 compounds included on the Annual Emissions Inventory for gas wells and related facilities. [16] On Table 3, we show the tons/year of seven compounds emitted from well pads, processing plants, and compressor stations in Washington County, PA. These seven compounds are shown because they are consistently emitted in larger mass than other compounds.

Household-level exposure to emissions

Household level exposure in this sample varied by year and location of the source relative to the residence. Table 4 shows the annual median and upper and lower limits of ambient air levels of emissions of the group of compounds included in this analysis: nitrogen oxides, carbon monoxide, VOCs, PM 2.5, and formaldehyde. These five compounds had the highest reported mass and known health effects.

Relationships between exposure and respiratory health outcomes

Although all reported one or more symptoms that began or worsened after the onset of drilling activity and could not be plausibly attributed to pre-existing or current medical conditions, or practices such as smoking, 28% of the sample did not report any respiratory symptoms at all. At least one respiratory symptom (i.e., “any respiratory symptom”) was reported by 72%; sore throat by 40%; cough and shortness of breath by 36% each; sinus problems by 26%, and wheezing by 16%. The majority (77%) lived within two kilometers of at least one source: 29% lived within 2 km of 1-9 sources; 25% within 2 km of 10-19 sources; and 23% within 2 km of 20 or more. The number of sources within 2 km ranged from 0-40.

Any respiratory symptom and exposure

For any respiratory symptom (i.e., at least one respiratory symptom reported), the overall model with 11 predictors was significant: $\chi^2(11) = 22.32$, $p = .022$ (Cox & Snell $r^2 = .231$; Nagelkerke $r^2 = .329$). As well, the Hosmer-Lemeshow test was not significant: $\chi^2(7) = 4.09$, $p = .769$. In regards to classification, sensitivity was 91.7% and specificity was 36%. As shown on Table 5, the following predictors are significant:

(1) South Sources: $b = -.464$, $p = .018$ (OR = 0.629, 95% CI = 0.428, 0.923) indicating that the higher the value for south source the lower the probability of having any symptom; and (2) West Sources: $b = .41$, $p = .01$ (OR = 1.51, 95% CI = 1.10, 2.06) indicating that the higher the value for west source the higher the probability of having any symptom.

Cough and exposure

For cough, the overall model with 11 predictors was significant: $\chi^2(11) = 27.06$, $p = .005$ (Cox & Snell $r^2 = .273$; Nagelkerke $r^2 = .373$). As well, the Hosmer-Lemeshow test was not significant: $\chi^2(7) = 8.03$, $p = .33$. In regards to classification, sensitivity was 67.7% and specificity was 74.1%. As shown in Table 6, the following predictors are significant: (1) West Sources: $b = .209$, $p = .04$ (OR = 1.23, 95% CI = 1.01, 1.51) indicating that the higher the value for west source the higher the probability of having the cough symptom and (2) water source: $b = 2.22$, $p = .001$ (OR = 9.19, 95% CI = 2.40, 35.13) indicating that those who have a well or other non- municipal water source have a higher probability of having the cough symptom than those who have a municipal water source.

Shortness of breath and exposure

For shortness of breath, the overall model with 11 predictors was significant: $\chi^2(11) = 25.54$, $p = .008$ (Cox & Snell $r^2 = .26$; Nagelkerke $r^2 = .355$). As well, the Hosmer-Lemeshow test was not significant: $\chi^2(7) = 1.73$, $p = .973$. In regards to classification, sensitivity was 45.2% and specificity was 87%. As shown on Table 7, the following predictors are significant: (1) South Sources: *b* = -.372, *p* = .049 (OR = 0.689, 95% CI = 0.476, 0.999) indicating that the higher the value for south source the lower the probability of having the shortness of breath symptom and (2) West Sources: *b* = .439, *p* = .003 (OR = 1.55, 95% CI = 1.16, 2.08) indicating that the higher the value for west source the higher the probability of having the shortness of breath symptom. There were no significant associations between the exposure measures and sore throat, sinus problems, or wheezing.

The results of our analysis suggest that an exposure metric including the number of sources in combination with wind direction may be a better predictor of new onset respiratory symptoms than the number of sources and wind direction combined with the reported emissions from those sources.

There are several possible explanations for our findings. Any respiratory symptom, cough, and shortness of breath were associated with the cardinal direction of the emission source. The levels of contamination in the air move along with weather systems. In this county, that is primarily from the southwest to the northeast. The periods of stability in a weather system vary (i.e., there are periods of stability and low dilution of pollutants). It is possible that weather systems that move in other directions, although they also carry pollutants, have differing periods of high stability and low dilution.

Our calculations of the annual hourly emission rate from a source assume a consistent rate and may not be reflective of the occurrence of peak emissions. Peak emissions may be more important in precipitating acute respiratory symptoms than average emissions.

Proximity to wells inherently captures peak events. In a similar vein, emission data do not capture exposures to flaring, impoundments or indoor off-gassing from contaminated water, all of which might result in peak exposures precipitating respiratory symptoms. Since these sources are typically co-located with wells, proximity to sources would better reflect these exposures.

Although four of the five compounds included in this analysis have recognized acute respiratory effects (NO_x, PM2.5, VOC's and formaldehyde), there may be other emissions with potent respiratory effects that were not included in the analysis. Future analyses could be limited to those emissions with established respiratory actions. Additionally, our results support the potential exposure presented by ground water. Ground water contamination associated with the gas wells may contribute to respiratory symptoms such as cough through off gassing during indoor use.

Our study used a convenience sample whose self-reported date of respiratory symptom onset fell within the year of the exposure estimate. The lack of precision in the temporal relationship to the exposure is a limitation of this study. However, self-reported symptoms were reviewed by a nurse practitioner, in a standardized clinical interview, and all records were reviewed to include only those symptoms that could not plausibly be explained by a co-occurring medical condition or a habit such as smoking. A further limitation is our focus on respiratory symptoms, when there are other health symptoms that have been associated with UNGD that we did not include.

Other approaches to estimation of exposures have considered distance from nearest wells and the number of wells. Our approach considered the intensity of the emissions;

temporal dilution factors; number of sources within a specified distance; and cardinal direction of those sources relative to the residence. The approach included not just well pads, but also processing plants and compressor stations. We feel that inclusion of infrastructure such as processing plants and compressor stations and the addition of cardinal direction provides a more precise assessment of exposure at the residence. The processes needed to consider cardinal directions are not overly cumbersome. When using approaches such as the inverse ratio of the distance squared, one can simply consider the direction from the source and flow of weather patterns in the area.

the direction from the source and flow of weather patterns in the area.

Conclusion

To our knowledge this is only the second study that included multiple sources of pollution (i.e., well pads, processing plants, and compressor stations) and the first study to incorporate weather and atmospheric conditions in the exposure estimate. Estimates of exposure typically characterize the sources within a specific radius of a residences. We suggest that future characterizations should consider the cardinal direction of the source from the residence. The impact of ambient air levels is unclear should be investigated in future studies.

Acknowledgements

This work was supported by funding from Heinz Endowments, Pittsburgh, PA.

References

[1] Schraufnagel, D.E.; Balmes, J.R.; Cowl, C.T.; De Matteis, S.; Jung, S-H.; Mortimer, K.; Perez-Padilla, R.; Rice, M.B.; Riojas-Rodriquez, H.; Sood, A.; Thurston, G.D.; To, T.; Vanker, N.; Wuebbles, E.J. Air pollution and noncommunicable diseases: A review by the Forum of International Respiratory Societies' Environmental Committee, Part 1: The damaging effects of air pollution. Chest. **2019,** *155*, 409–16.

[2] Schraufnagel, D.E.; Balmes, J.R.; Cowl, C.T.; De Matteis, S.; Jung, S-H.; Mortimer, K.; Perez-Padilla, R.; Rice, M.B.; Riojas-Rodriquez, H.; Sood, A; Thurston, G.D.; To, T.; Vanker, N.; Wuebbles, E.J. Air pollution and noncommunicable diseases: A review by the Forum of International Respiratory Societies' Environmental Committee, Part 2: Air pollution and organ systems. Chest. **2019,** *155*, 417–26.

[3] McKenzie, L.M.; Guo, R.; Witter, R.Z.; Savitz, D.A.; Newman, L.S.; Adgate, J.L. Birth outcomes and maternal residential proximity to natural gas development in rural Colorado. Environ. Health Perspect. **2014,** *122*, 412–7.

[4] Stacy, S.L.; Brink, L.L.; Larkin, J.C.; Sadovsky, Y.; Goldstein, B.D.; Pitt, B.R., Talbott, E.O. Perinatal outcomes and unconventional natural gas operations in Southwest Pennsylvania. PloS One. **2015,** *10*, e0126425.

[5] Casey, J.A.; Savitz, D.A.; Rasmussen, S.G.; Ogburn, E.L.; Pollak, J.; Mercer, D.G., Schwartz, B. Unconventional natural gas development and birth outcomes in Pennsylvania, USA. Epidemiol. **2016,** *27*, 163–72.

[6] McKenzie, L.M.; Allshouse, W.B.; Byers, T.E.; Bedrick, E.J.; Serdar, B.; Adgate, J.L. Childhood hematologic cancer and residential proximity to oil and gas development. PloS One. **2017,** *12*, e0170423.

[7] Tustin, A.W.; Hirsch, A.G.; Rasmussen, S.G.; Casey, J.A.; Bandeen-Roche, K.; Schwartz, B.S. Associations between unconventional natural gas development and nasal and sinus, migraine headache, and fatigue symptoms in Pennsylvania. Environ. Health Perspect. **2016,** *125*, 189-197.

[8] Graham, I.J.; Tang, X.; Sellers, S.; Crisp, J.; Horwitz, D.; Muehlenbachs, L.; Krupnick, A.; Carey, D. Increased traffic accident rates associated with shale gas drilling in Pennsylvania. Accid. Anal. Prev. **2015,** *74*, 203–9.

[9] Casey, J.A.; Wilcox, H.C.; Hirsch, A.G.; Pollak, J.; Schwartz, B.S. Associations of unconventional natural gas development with depression symptoms and disordered sleep in Pennsylvania. Sci. Rep. **2018,** *8*, 11375.

[10] Rasmussen, S.G.; Ogburn, E.L.; McCormack, M.; Casey, J.A.; Bandeen-Roche, K.; Mercer, D.G.; Schwartz, B.S. Association between unconventional natural gas development in the Marcellus Shale and asthma exacerbations. JAMA Intern. Med. **2016,** *176*, 1334-43.

[11] Rabinowitz, P.M.; Slizovskiy, I.B.; Lamers, V.; Trufan, S.J.; Holford, T.R.; Dziura, J.D.; Peduzzi, P.N.; Kane, M.J.; Reif, J.S.; Weiss, T.R.; Stowe, M.H. Proximity to natural gas wells and reported health status: Results of a household survey in Washington County, Pennsylvania. Environ. Health Perspect. **2015,** *123*, 21–6.

[12] Weinberger, B.; Greiner, L.H.; Walleigh, L.; Brown, D. Health symptoms in residents living near shale gas activity: A retrospective record Review from the environmental health project. Prev. Med. Rep. **2017,** *8*, 112–5.

[13] Hill, E.L. Shale gas development and infant health: Evidence from Pennsylvania. J. Health Econ. **2018,** *61*, 134–50.

[14] Koehler, K.; Ellis, J.H.; Casey, J.A.; Manthos, D.; Bandeen-Roche, K.; Platt, R.; Schwartz, B.R. Exposure assessment using secondary data sources in unconventional natural gas development and health studies. Environ. Sci. Technol. **2018,** *52*, 6061–9.

[15] Pennsylvania Department of Environmental Protection. Air emissions inventory for the natural gas industry fact sheet [Internet]. [cited 2019 Mar 25]. Available from: http://files.dep.state.pa.us/Air/AirQuality/AQPortalFiles/Regulations%20and%20Clean%20Air%20Plans/Natural_Gas_Inventory_Fact_Sheet_02-11-13.pdf

[16] Air Emissions Report - Power BI Report Server [Internet]. [cited 2019 Mar 17]. Available from: http://www.depgreenport.state.pa.us/powerbiproxy/powerbi/Public/DEP/AQ/PBI/Air_Emissions_Report

[17] Brown, D.R.; Lewis, C.; Weinberger, B.I. Human exposure to unconventional natural gas development: A public health demonstration of periodic high exposure to chemical mixtures in ambient air. J. Environ. Sci. Health Part A Tox. Hazard Subst. Environ. Eng. **2015,** *50*, 460–72.

[18] LatLong.net [Internet]. [cited 2019 Aug 3]. Available from: https://www.latlong.net/

[19] Pasquill, F. *Atmospheric Diffusion: The Dispersion of Windborne Material from Industrial and other Sources*; D. Van Norstrand: London, 1962.

[20] Burton, R. Atmospheric dispersion [Internet]. [cited 2019 Mar 24]. Available from: http://homepages.see.leeds.ac.uk/~lecrrb/dispersion/notes.pdf

[21] Brown, D.; Weinberger, B.; Lewis, C.; Bonaparte, H. Understanding exposure from natural gas drilling puts current air standards to the test. Rev. Environ. Health. **2014,** *29*, 277–92.

[22] NOAA National Centers for Environmental Information. Data tools: Local climatological data [Internet]. [cited 2019 Mar 24]. Available from: https://www.ncdc.noaa.gov/cdo-web/datatools/lcd

[23] Menard, S. *Logistic Regression*; Sage: Los Angeles, CA, 2010.

[24] Osborne, J.W. *Best Practices in Logistic Regression*; Sage: Los Angeles, CA, 2015.

FIGURE CAPTIONS

Figure 1. Vertical mixing by stability class and distance from source [19]

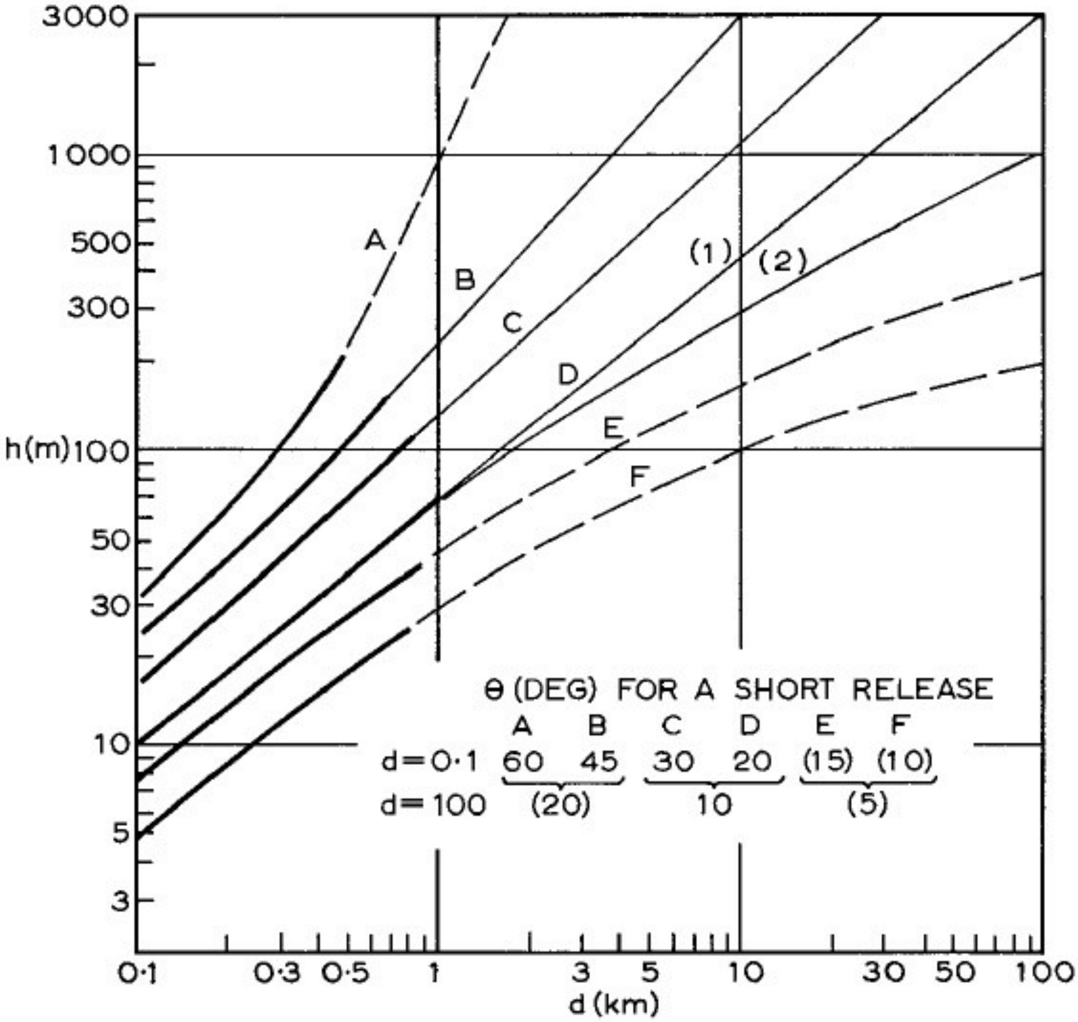
3000
1000
500
300
h(m)100
50
30
10
5
3
0·1
0·3
0·5
1
3
5
10
30
50
100
d (km)
A
B
C
D
E
F
(1)
(2)
θ (DEG) FOR A SHORT RELEASE
A B C D E F
d=0·1 60 45 30 20 (15) (10)
d=100 (20) 10 (5)

Fig. 1

Table 1. Air stability classes [19]

Wind Speed (mph)	Day Clear or just a few clouds	Day < 50% cloud cover	Day > 50% cloud cover	Day > 80% cloud cover	Night < 50% cloud cover	Night > 50% cloud cover
<5	A1	AB6	B11	D16	F26	E21
5-7	AB2	B7	C12	D17	F27	E22
7-11	B3	BC8	C13	D18	E28	D23
11-13	C4	CD9	D14	D19	D29	D24
>13	C5	D10	D15	D20	D30	D25

Table 2. Ambient air levels in ug/m^3 from a hypothetical source emitting 300 grams per hour by stability class

	Distance from Source						
Stability Class	≤0.1 km	>0.1 km, ≤0.5 km	>0.5 km, ≤1 km	>1 km, ≤2 km	>2km, ≤3 km	>3km, ≤5 km	>5km, ≤10 km
A1	175	5	1	<1	<1	<1	0
AB2	125	8	2	<1	<1	<1	0
B3	75	8	2	<1	<1	<1	0
C4	125	22	7	2	1	<1	<1
C5	100	19	6	1	<1	<1	<1
AB6	175	12	3	<1	<1	<1	0
B7	150	13	4	<1	<1	<1	0
BC8	125	17	6	1	<1	<1	<1
CD9	175	28	10	3	1	<1	<1
D10	200	29	10	4	2	<1	<1
B11	200	18	6	1	<1	<1	<1
C12	250	41	14	4	1	<1	<1
C13	150	25	9	2	1	<1	<1
D14	225	33	12	4	2	<1	<1
D15	200	29	10	4	2	<1	<1
D16	625	87	32	12	6	2	<1
D17	450	62	23	9	4	1	<1
D18	275	39	14	5	2	1	<1
D19	225	33	12	4	2	<1	<1
D20	200	29	10	4	2	<1	<1
E21	875	150	73	35	21	10	3

E22	625	100	52	25	15	7	2
D23	275	39	14	5	3	1	<1
D24	225	33	12	4	2	<1	<1
D25	200	29	10	4	2	<1	<1
F26	1400	225	100	56	33	15	5
F27	1000	150	83	40	23	11	4
E28	375	78	33	16	9	4	1
D29	225	33	12	4	2	<1	<1
D30	200	29	10	4	2	<1	<1

Table 3. Tons per year of carbon dioxide, methane, nitrogen oxides, carbon monoxide, volatile organic compounds (VOCs), PM2.5, and formaldehyde emitted in Washington County PA 2012-2015

Compound	2012	2013	2014	2015	2016
Carbon dioxide	681,402	731,097	897,398	1,028,463	1,293,388
Methane	9,204	7,217	3,069	7,684	8,276
Nitrogen oxides	2,097	2,875	3,931	4,546	3,544
Carbon monoxide	939	939	1,375	2,087	1,742
VOCs	639	1,658	1,548	1,932	2,078
PM2.5	66	102	157	111	85
Formaldehyde	54	55	55	61	63

Table 4. Median ambient air levels of emissions of nitrogen oxides, carbon monoxide, VOCs, PM 2.5, and formaldehyde by year

Year	North	East	South	West
2012	0	0	0	25
	(0,1562.16)	(0,999.36)	(0,3972.58)	(0,1556.67)
2013	93.51	145.2	206.6	268.0
	(0,945.5)	(0,1208.8)	(0,6638.3)	(0,761.8)
2014	0	7.96	0	5.80
	(0,1301.24)	(0,3619.44)	(0,3902.76)	(0,8670.90)
2015	197.03	0.09	238.44	307.16
	(0,3784.57)	(0,4542.18)	(0,2684.26)	(0,2169.16)
2016	312.9	73.00	327.4	460.0
	(0,6065.00)	(0,1027.0)	(0,3661.6)	(0,3057.0)

Table 5. Associations between exposure measures and any reported respiratory symptom

	B	S.E.	p-value	OR	CI Lower	CI Upper
NorthEM	0	0	0.795	1	0.999	1.001
EastEM	0	0	0.263	1	0.999	1
SouthEM	0.001	0.001	0.147	1.001	1	1.002
WestEM	0	0	0.202	1	0.999	1
NorthSources	-0.259	0.147	0.078	0.772	0.578	1.03
SouthSources	-0.464	0.196	0.018	0.629	0.428	0.923
WestSources	0.411	0.159	0.01	1.508	1.104	2.06
EastSources	0.304	0.183	0.097	1.355	0.947	1.94
Water Source	0.177	0.694	0.799	1.193	0.306	4.649
Gender	-0.119	0.622	0.848	0.888	0.262	3.007
Age	0.031	0.031	0.316	1.032	0.971	1.097
Constant	-0.673					

Note: EM=emissions; B = unstandardized logit coefficient; S.E. = standard error; OR = odds ratio; CI = 95% confidence interval

Table 6. Associations between exposure measures and cough

	B	S.E.	p-value	OR	CI Lower	CI Upper
NorthEM	0	0	0.957	1	0.999	1.001
EastEM	-0.001	0.001	0.084	0.999	0.998	1
SouthEM	0	0	0.4	1	1	1.001
WestEM	-0.001	0	0.153	0.999	0.998	1
NorthSources	-0.026	0.107	0.804	0.974	0.79	1.201
SouthSources	-0.267	0.146	0.068	0.766	0.575	1.02
WestSources	0.209	0.102	0.04	1.232	1.009	1.505
EastSources	0.086	0.106	0.414	1.09	0.886	1.341
Water Source	2.218	0.684	0.001	9.186	2.403	35.125
Gender	0.404	0.618	0.514	1.497	0.446	5.029
Age	0.01	0.028	0.724	1.01	0.957	1.066
Constant	-2.528					

Note: EM=emissions; B = unstandardized logit coefficient; S.E. = standard error; OR = odds ratio; CI = 95% confidence interval

Table 7. Associations between exposure measures and shortness of breath

	B	S.E.	p-value	OR	CI Lower	CI Upper
NorthEM	0	0	0.546	1	1	1.001
EastEM	-0.001	0	0.15	0.999	0.999	1
SouthEM	0.001	0.001	0.164	1.001	1	1.002
WestEM	-0.002	0.001	0.078	0.998	0.996	1
NorthSources	-0.142	0.124	0.254	0.868	0.68	1.107
SouthSources	-0.372	0.189	0.049	0.689	0.476	0.999
WestSources	0.439	0.15	0.003	1.551	1.156	2.08
EastSources	0.259	0.156	0.097	1.295	0.954	1.759
Water Source	0.261	0.651	0.689	1.298	0.362	4.654
Gender	-1.027	0.586	0.08	0.358	0.114	1.129
Age	0.019	0.029	0.51	1.019	0.964	1.078
Constant	-1.361					

Note: EM=emissions; B = unstandardized logit coefficient; S.E. = standard error; OR = odds ratio; CI = 95% confidence interval