



Article

# Spatiotemporal Correlation Analysis of Hydraulic Fracturing and Stroke in the United States

Chuanbo Hu <sup>1</sup>, Bin Liu <sup>2</sup>, Shuo Wang <sup>3</sup> , Zhenduo Zhu <sup>4</sup> , Amelia Adcock <sup>5</sup> , James Simpkins <sup>6</sup> and Xin Li <sup>1,\*</sup>

<sup>1</sup> Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26505, USA

<sup>2</sup> Department of Management Information Systems, West Virginia University, Morgantown, WV 26505, USA

<sup>3</sup> Department of Radiology, Washington University, St. Louis, MO 63110, USA

<sup>4</sup> Department of Civil, Structural and Environmental Engineering, University at Buffalo, Buffalo, NY 14260, USA

<sup>5</sup> Department of Neurology, West Virginia University, Morgantown, WV 26505, USA

<sup>6</sup> Department of Neuroscience, West Virginia University, Morgantown, WV 26505, USA

\* Correspondence: xin.li@mail.wvu.edu

**Abstract:** Hydraulic fracturing or fracking has led to a rapid growth of oil and gas production in the United States, but the impact of fracking on public health is an important but underresearched topic. We designed a methodology to study spatiotemporal correlations between the risk of fracking and stroke mortality. An annualized loss expectancy (ALE) model is applied to quantify the risk of fracking. The geographically and temporally weighted regression (GTWR) model is used to analyze spatiotemporal correlations of stroke mortality, fracking ALE, and nine other socioeconomic and health-related factors. The analysis shows that fracking ALE is moderately correlated with stroke mortality at ages over 65 in most states of fracking, in addition to cardiovascular disease and drug overdose being positively correlated with stroke mortality. Furthermore, the correlations between fracking ALE and stroke mortality in men appear to be higher than in women near the Marcellus Shale, including Ohio, Pennsylvania, West Virginia, and Virginia, while stroke mortality among women is concentrated in the Great Plains, including Montana, Wyoming, New Mexico, and Oklahoma. Lastly, within two kilometers of the fracking mining activity, the level of benzene in the air was found to be significantly correlated with the fracking activity in Colorado.

**Keywords:** fracking; stroke mortality; annualized loss expectancy (ALE); geographically and temporally weighted regression (GTWR); spatiotemporal analysis



**Citation:** Hu, C.; Liu, B.; Wang, S.; Zhu, Z.; Adcock, A.; Simpkins, J.; Li, X. Spatiotemporal Correlation Analysis of Hydraulic Fracturing and Stroke in the United States. *Int. J. Environ. Res. Public Health* **2022**, *19*, 10817. <https://doi.org/10.3390/ijerph191710817>

Academic Editors: Agata Stanek and Edyta Ewa Sutkowska

Received: 19 July 2022

Accepted: 26 August 2022

Published: 30 August 2022

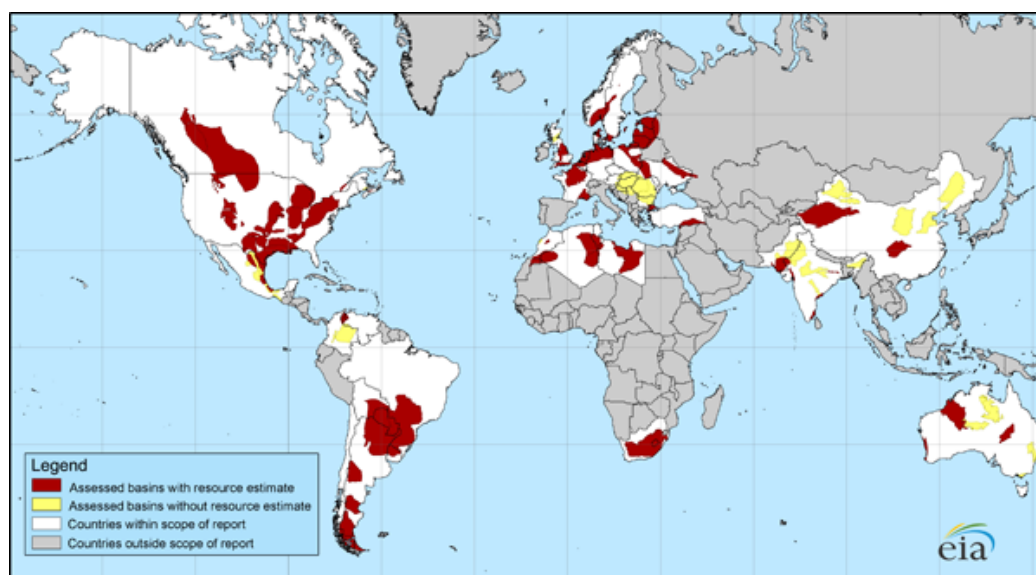
**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Hydraulic fracturing [1], also known as fracking, is a geochemical engineering process by which large volumes of water combined with chemical and sand proppants are injected into tight formations with high pressure to fracture and facilitate recovery of unconventional reserves of oil and gas [2]. With the development of fracking technology in the United States, shale gas is becoming an increasingly important source of natural gas, and interest has spread to countries around the world with shale gas storage. The US Energy Information Administration (EIA) assessed 137 shale formations in 42 countries around the world [3], and the distribution of assessed shale gas and shale oil basins of the world is shown in Figure 1. They represent 10% of the world's crude oil and 32% of the world's natural gas. However, as of 2013, only three countries (United States, Canada, and China) have significant commercial shale gas production due to technical limitations and local laws.



**Figure 1.** Mapping the distribution of assessed global shale gas and shale oil basins [3]. Source from United States Energy Information Administration (EIA) and United States Geological Survey.

Currently, the United States is the world's largest producer of both natural gas and crude oil. According to the US Energy Information Administration (EIA), there were approximately 23,000 fracking wells in the US in 2000. In 2015, the number of fracking wells increased rapidly to approximately 300,000, representing 67% of United States natural gas production and 51% of United States crude oil production [4]. Despite the economic benefits of fracking, this expansion has brought the industrial activities of oil and gas development closer to backyards and communities, increasing the risk of human exposure to new contaminants and threats [5,6].

The health implications and effects of fracking have not been adequately studied [7–12]. There is a growing body of research studies on the negative impact of fracking on air and water quality [13–17], as well as public health [18–22]. Although health discussions have focused on drinking water contamination, particularly in the eastern US, there is growing interest in studying a variety of health threats arising from air pollution [7,23]. Health threats from air pollution vary significantly across environments [24–27]. For example, research has attempted to link fracking pollution with unhealthy levels of smog and toxic air pollutants [28]. Exposure to this pollution can cause eye, nose, throat, respiratory disease, birth defects, cancer, or premature death [29]. However, little is known about whether fracking can cause life-threatening conditions such as stroke [8].

Stroke as a neurological disease is the leading cause of long-term adult disability and the fifth leading cause of death in the United States [30], with approximately 795,000 stroke events annually. Stroke belt refers to a consistent pattern of striking geographic variation in stroke mortality rates within the United States [31]. It covers 11 states in the southeast US with an unusually high incidence of stroke and other forms of cardiovascular disease. Factors that explain the prevalence of excessive stroke in the stroke belt include differences in socioeconomic status (e.g., employment rate and marital rate), risk factors (e.g., smoking and unhealthy diet), and prevalence of common chronic diseases (e.g., diabetes and heart disease) [32,33]. A recent study has shown that the highest contributors to the Stroke Belt include a higher burden of risk factors, higher levels of inflammation and infection, and lower socioeconomic status, while environmental exposures and lifestyle choices are considered lesser contributors [34].

*Does fracking induce a higher risk of stroke?* Although a potential connection between fracking and stroke has been mentioned in the literature (e.g., [35–37]), there is no systematic study to address this question. The closest research to this work is the study on the impact of fracking on water pollution [38] and air pollution [39], but its research data are limited

to the local area. Furthermore, the spatial extent of the public health impact of hydraulic fracturing is a question that existing research attempts to answer. For example, the distance to the nearest fracking well has been used as an important indicator to analyze the spatial correlation between fracking and infant health [40]. A spatial analysis method has been designed to quantify the environment at risk of Marcellus Shale fracking in the state of PA, USA [41]. The study [42] verified that people within 0.8 km of a fracking well are particularly at risk to their health. However, few studies have analyzed the spatiotemporal impacts of hydraulic fracturing on public health. The goal of this study is to address this question from a geographic information system (GIS) perspective using the extension of the geographically weighted regression (GWR) method [43]: geographically and temporally weighted regression (GTWR) [44]. We performed a detailed regression analysis of the stroke and fracking data using GTWR. We hope that this work can shed light on the relations between fracking and stroke risk and stimulate more quantitative studies on the health risk of fracking, which can better inform decision makers about energy and public health policy.

## 2. Methodology

### 2.1. Study Area and Data Collection

To study the spatiotemporal correlations of stroke mortality and fracking, the present study has chosen 49 states in the US as the study area. Alaska is not included due to its geographical isolation. Figure 2 shows the stroke death rate per 100,000 people over 65 years of age and all sites of fracking activity before 2018. To explore the impact of fracking on stroke mortality, we divide the 49 states into *fracking states* and *non-fracking states*. As a result, there are 24 fracking states that had fracking activities (including 19 states with active fracking and 5 states with little fracking) and 25 non-fracking states that did not have fracking activity by 2018.

Stroke is closely related to people's behavior habits (tobacco use, high cholesterol diet, and physical activity index), socioeconomic status (family mean income, marital rate, and employment rate), and other diseases (cardiovascular, overdose, and diabetes) [33,45]. Therefore, these variables have been selected for comparison with risk factors for fracking. Table 1 shows the details of the dependent and explanatory variables and their data sources. All data was collected in the US from 2010 to 2018. The scale of the data we collected is state-level except for fracking, because the county-level data contain a lot of missing data that can cause problems with the analysis. Furthermore, since stroke is a chronic disease, we processed the fracking data into the cumulative number of fracking wells. In other words, the number of fracking wells in any year includes all fracking wells before that year.

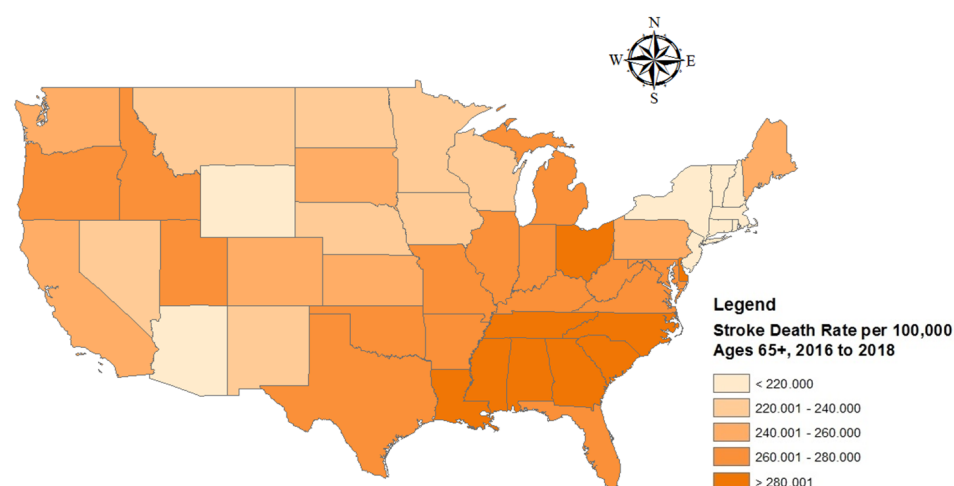
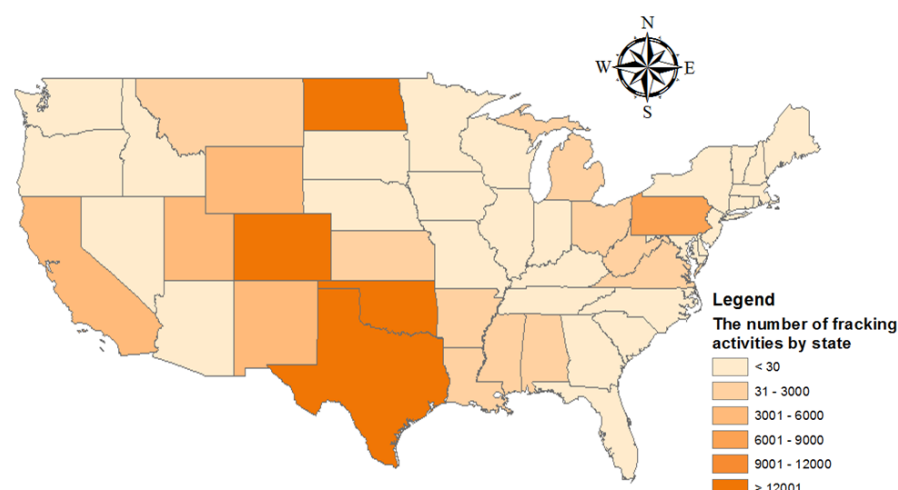


Figure 2. Cont.



**Figure 2.** Map of the study area and the distribution of the stroke death rate per 100,000 population 65 years and older (**upper**) and US fracking activities by state prior to 2018 (**lower**).

**Table 1.** Description of the data used in this study (CDC: Centers for Disease Control and Prevention; USCB: The United States Census Bureau; FF: United States FracFocus; HAPs: Hazardous Air Pollutants; EPA: The United States Environmental Protection Agency).

Type	Variables	Description	Source
Fracking	Fracking activity	The location of Fracking wells	FF
Stroke mortality	65+ stroke mortality	Stroke deaths rate over 65 per 100,000	CDC
	Male stroke mortality	Male Stroke deaths rate per 100,000	CDC
	Female stroke mortality	Female Stroke deaths rate per 100,000	CDC
Disease	Diabetes	Proportion of diagnosed diabetes	CDC
	Cardiovascular	Cardiovascular deaths rate over 65 per 100,000	CDC
	Overdose	Drug overdose death rates	CDC
	Hypertension	High blood pressure deaths over 65 per 100,000	CDC
	Obesity	Adults with a BMI > 30	CDC
Behavior	Tobacco use	Current cigarette use by adults	CDC
	High cholesterol	High total cholesterol among adults	CDC
	Physical activity index	Physical Inactivity Prevalence	CDC
	Heavy Drink	8 or more drinks per week (female) or 15 or more drinks per week (male)	CDC
Socioeconomic	Mean income	Family income by number of workers	USCB
	Marital rate	Proportion of married population	USCB
	Employment rate	Proportion of employed population	USCB
	Education	Bachelor's degree or higher	USCB
HAPs	Butadiene	Concentration monitoring data for Butadiene	EPA
	Benzene	Concentration monitoring data for Benzene	EPA
	Formaldehyde	Concentration monitoring data for Formaldehyde	EPA
	Acetaldehyde	Concentration monitoring data for Acetaldehyde	EPA

## 2.2. Quantitative Risk Analysis: Annual Loss Expectancy of Fracking

Recent research illustrated that 2–16% of oil and gas wells spill liquids every year [46]. These accidents have caused fracking chemicals to contaminate drinking water and air, further raising serious public health risks [47–49]. To quantify the public health risk of fracking, the annualized loss expectancy (ALE) model [50] has been applied to assess the public health-related ALE caused by fracking by state, which is abbreviated as fracking ALE.

ALE is the product of the annual rate of occurrence (ARO) and the single loss expectancy (SLE) [51] caused by fracking, as shown in Equation (1).

$$ALE = ARO \times SLE \quad (1)$$

where fracking ARO is the annual occurrence of pollutant leakage due to fracking, which is equal to the Fracking Density per square kilometer (FD) multiplied by the annual rate of occurrence per fracking per square kilometer (AROF),  $ARO = FD \times AROF$ .

Fracking SLE is defined as a population with negative health impacts expected from the occurrence of fracking accidents. For risk calculation, SLE is used to calculate a single loss when a specific event occurs. Fracking SLE is calculated by multiplying the Population Density per square kilometer (PD) by the fracking exposure factor (EF),  $SLE = PD \times EF$ , as the higher the population density in fracking-active states, the greater the negative impact of fracking on public health.

Fracking ALE represents the product of AROF, FD, PD, and EF, which means the average loss per year of environmental pollution caused by fracking on public health. The FD and PD of the  $j$ th state in the  $i$ th year are  $FD_{ij}$  and  $PD_{ij}$ , respectively. As an important explanatory variable for the regression in Section 2.3.3, ALE is normalized as follows:

$$Normalized\ ALE_{ij} = \frac{ALE_{ij} - Min(ALE)}{Max(ALE) - Min(ALE)} \quad (2)$$

where  $Normalized\ ALE_{ij}$  is the normalized ALE in state  $j$  in year  $i$ . Since AROF and EF are constants, Equation (2) can be further formulated as

$$Normalized\ ALE_{ij} = \frac{FD_{ij} \times PD_{ij} - Min(FD \times PD)}{Max(FD \times PD) - Min(FD \times PD)} \quad (3)$$

### 2.3. Regression Model

#### 2.3.1. Multicollinearity

We performed multicollinearity diagnostics before applying the regression model, since several explanatory variables could be highly correlated. Multicollinearity means that there is a high linear correlation between several specific explanatory variables, which could lead to bias in explaining the significance and associations of other variables. We adopted the variance inflation factor (VIF) [52], a metric of the severity of multicollinearity, to eliminate this problem. Explanatory variables with a VIF greater than 10 are considered to cause multicollinearity and should be excluded from the model [53].

#### 2.3.2. Spatial Autocorrelation

As a commonly used spatial autocorrelation test, Moran's I test represents the spatial autocorrelation of a single explanatory variable and can be expressed as [54].

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

Moran's I ranges from  $-1$  to  $1$ . A higher positive value indicates that closer observations have more similar attribute values, whereas farther observations have more distinct attribute values, indicating spatial aggregation. Negative values represent a spatially distributed distribution and a zero value represents a spatially random distribution. The null hypothesis of the Moran's I test indicates that the explanatory variables are spatially independent. It indicates that the Moran's I is close to zero. The Z-score is used as a significant indicator to measure Moran's I to verify the null hypothesis, whose formula is as follows [54].

$$Z(I) = \frac{I - E(I)}{\sqrt{Var(I)}} \quad (5)$$



where  $E(I)$  and  $Var(I)$  are the expectation and standard deviation of Moran's I, respectively. The significance level in this study is established as the  $p$ -value  $< 0.05$ .

### 2.3.3. Geographical and Temporal Weighted Regression (GTWR)

Spatiotemporal data analysis provides a series of important tools to solve problems such as correlation analysis of spatiotemporal data, spatiotemporal pattern analysis, and spatiotemporal prediction problems [44,55–59]. To analyze the spatio-temporal correlation between risk factors (see Table 1) and stroke mortality, the GTWR model was selected as the regression model. Compared to traditional Geographically Weighted Regression (GWR) [43], which only considers spatial features, the GTWR model [44] considers the non-stationary effect in space and time. Therefore, it was adopted to explore the spatiotemporal heterogeneity of the influence of fracking on different stroke mortality (i.e., age-based and gender-based) under the constraint of spatiotemporal differences. By establishing a three-dimensional elliptical coordinate system (including time, longitude, and latitude) in which the temporal dimension is the vertical dimension, in addition to the two horizontal spatial dimensions (longitude and latitude), the model can describe the spatio-temporal influence via the regression coefficients corresponding to explanatory variables. The GTWR model is described as

$$Y_i = \beta_0(u_i, v_i, t_i) + \sum_k \beta_k(u_i, v_i, t_i) X_{ik} + \varepsilon_i \quad (6)$$

where  $u$  and  $v$  represent longitude and latitude, respectively.  $\beta_0(u_i, v_i, t_i)$  represents the intercept item of state with centroid at  $(u_i, v_i)$  in year  $t_i$ ;  $k$  is the number of explanatory variables;  $\beta_k(u_i, v_i, t_i)$  is the regression coefficient of the  $k$ th explanatory variable in year  $t_i$ ;  $X_{ik}$  is the  $k$ th explanatory variable. These explanatory variables are defined in Table 1.

The correctional values of the Akaike Information Criterion (AICc) [60] is an important metric and often used to select explanatory variables and determine the final model with the lowest AICc. In this study, the ArcGIS GTWR plugin was used to analyze the spatio-temporal correlation between stroke mortality and risk factors such as fracking ALE.

## 3. Results

### 3.1. Model Comparison

To solve the multicollinearity problem between multiple variables, the variables with a VIF greater than 10 (i.e., hypertension and obesity) were removed. The results of the VIF values of significant explanatory variables are given in Table A1 for stroke mortality over 65 years and Table A2 stroke mortality 45–64 in Appendix A. Additionally, Moran's I statistics were calculated to determine whether the explanatory variables in Tables A1 and A2 are spatially associated. The results of the Moran's I test are given in Table A3 in Appendix A. The 10 selected variables with a  $p$ -value below 0.05 were included in the regression model, indicating that all variables are spatially autocorrelated.

Furthermore, to increase the significance of the regression variable, two explanatory variables (i.e., heavy drink and education) have been removed using stepwise selection based on AICc. Furthermore, a comparison with three baseline models was implemented, including Ordinary Least Squares (OLS) [61], Temporally Weighted Regression (TWR) [44] and GWR [43]) to evaluate the performance of the GTWR model [44]. As shown in Table 2. GTWR outperforms OLS, TWR, and GWR in model fitting, demonstrating that it better explains dependent variable stroke mortality. Taking the fracking state model as an example, the values of  $R^2$  increase from 0.757 in the OLS model, 0.768 in the TWR model, and 0.933 in the GWR model, to 0.970 in the GTWR model. The AICc reduces from  $-401.129$  in the OLS,  $-399.780$  in the TWR model and  $-534.502$  in the GWR model, to  $-564.090$  in the GTWR model (the lower, the better). The explanatory power increases significantly as spatial information and temporal information are considered in the model. In the rest of the paper, we will only analyze the results of the GTWR model.

**Table 2.** Comparison results of OLS, TWR, GWR, and GTWR models.

	Fracking States			Non-Fracking States		
	AICc	R <sup>2</sup>	Adjusted R <sup>2</sup>	AICc	R <sup>2</sup>	Adjusted R <sup>2</sup>
OLS [61]	−401.129	0.757	0.745	−293.313	0.601	0.584
TWR [44]	−399.780	0.768	0.757	−318.285	0.691	0.678
GWR [43]	−534.502	0.933	0.929	−481.016	0.897	0.892
GTWR [44]	<b>−564.090</b>	<b>0.970</b>	<b>0.968</b>	<b>−487.886</b>	<b>0.931</b>	<b>0.928</b>

Estimates of regression coefficients for 65+ stroke mortality in fracking states and non-fracking states are given in Tables 3 and 4. The results of both models showed the positive effect of cardiovascular and overdose on 65+ stroke mortality, and the negative effect of marital rate and employment rate on 65+ stroke mortality. These results share some similarities with other findings of previous work [62–64]. Furthermore, Table 3 shows that most fracking ALE coefficients are positively associated with stroke mortality at 65 years, although the correlation was much lower than for stroke mortality variables such as cardiovascular disease and overdose. Tables 3 and 4 show that there is no positive correlation between high cholesterol and stroke mortality. The result shares some similarities with other observations from previous work [33]. Furthermore, there is no correlation between tobacco use and 65+ stroke mortality, as its  $p$ -value  $> 0.1$  (see Table A1).

**Table 3.** Estimation of the GTWR model for 65+ stroke mortality in fracking states.

Variables	MIN	LQ	MED	UQ	MAX	AVG
<b>Related-disease risk factors</b>						
Diabetes	−0.661	−0.316	−0.001	0.218	0.632	−0.018
Cardiovascular	−1.506	−0.070	0.269	0.524	1.141	0.232
Overdose	−0.861	−0.030	0.389	0.693	1.148	0.347
<b>Behavior risk factors</b>						
Tobacco use	−0.973	−0.342	−0.157	−0.035	0.351	−0.197
High cholesterol	−1.346	−0.325	−0.149	0.090	0.415	−0.139
PAI	−0.585	−0.161	−0.080	0.020	0.521	−0.056
<b>Socioeconomic risk factors</b>						
Mean income	−1.442	−0.215	0.147	0.367	0.623	0.039
Marital rate	−1.459	−0.984	−0.677	−0.409	0.720	−0.663
Employment rate	−2.106	−0.604	−0.014	0.128	0.627	−0.223
<b>Fracking risk factor</b>						
Fracking ALE	−0.327	0.041	0.116	0.154	0.394	0.094
Intercept	−0.009	0.487	0.669	0.792	3.747	0.715

**Table 4.** Estimation of the GTWR model for 65+ stroke mortality in non-fracking states.

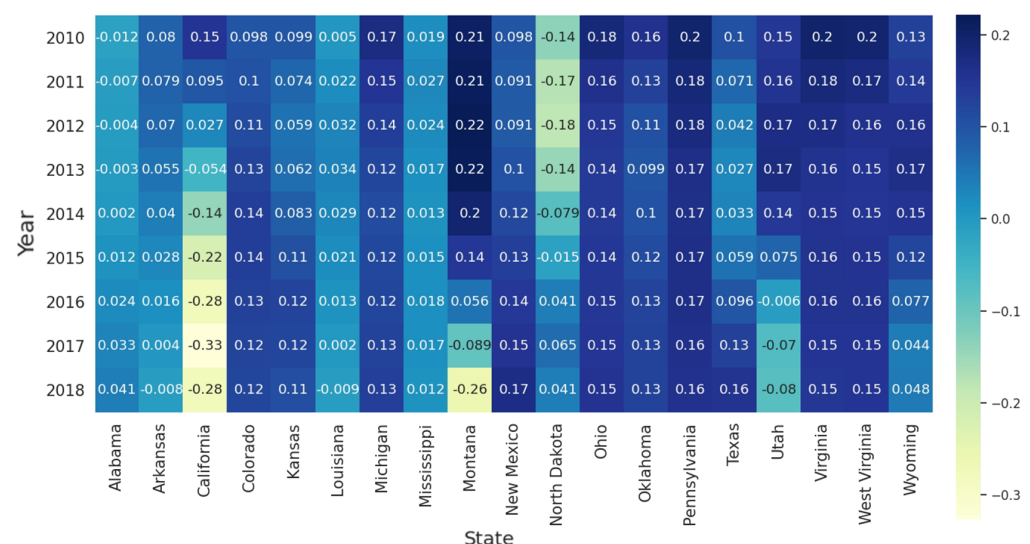
Variables	MIN	LQ	MED	UQ	MAX	AVG
<b>Related-disease risk factors</b>						
Diabetes	−0.603	−0.085	0.227	0.314	0.467	0.109
Cardiovascular	−0.354	0.272	0.357	0.427	0.896	0.305
Overdose	−1.341	−0.189	0.166	0.286	0.805	0.037
<b>Behavior risk factors</b>						
Tobacco use	−0.577	−0.036	0.102	0.222	0.71	0.08
High cholesterol	−0.841	−0.539	−0.296	−0.165	0.175	−0.333
PAI	−0.493	−0.101	0.043	0.233	0.451	0.063
<b>Socioeconomic risk factors</b>						
Mean income	−2.403	−0.684	−0.412	−0.222	0.476	−0.479
Marital rate	−1.448	−0.651	−0.078	0.265	0.354	−0.221
Employment rate	−1.657	−1.07	−0.615	−0.144	1.122	−0.507

### 3.2. Spatiotemporal Features of Fracking ALE Coefficients

We analyze the temporal and spatial characteristics of fracking ALE using the average values of the regression coefficients, which help to explore the temporal trends and spatial differences of fracking ALE on stroke mortality at ages 65 and older.

#### 3.2.1. Temporal Features of Fracking ALE Coefficients

The aforementioned improvement of GTWR is extended by incorporating the temporal dimension into the traditional GWR model. From the results of the GTWR model, we can obtain the time series of the yearly fracking ALE coefficients. Figure 3 presents the fluctuation of the average coefficients of the Fracking ALE variables over a 9-year period (from 2010 to 2018). Negative coefficients indicate the reverse correlation between the dependent and explanatory variables, and vice versa. 19 states with active fracking activity were discussed and 5 states with very few fracking activities were removed [65]. As Figure 3 shows, the positive correlation between fracking ALE and stroke mortality (65+) in California, Utah, Alabama, Louisiana, and Oklahoma decreased significantly year by year. In contrast, the positive correlation between fracking ALE and stroke mortality (65+) in North Dakota increased significantly. The coefficient of Fracking ALE in West Virginia, Ohio, Pennsylvania, Michigan, and Virginia slowly decreases. The ALE coefficient for fracking in Colorado and New Mexico first increases and then decreases.



**Figure 3.** Temporal distribution of the average coefficients of the ALE of Fracking for 19 states with active fracking.

#### 3.2.2. Spatial Features of Fracking ALE Coefficients

An important feature of the GTWR model is that the estimated coefficients are mappable for visual analysis. The spatial distributions of the effects of explanatory variables on ALE fracking are visualized in Figure 4. This study sets zero as a threshold to distinguish positive and negative effects. Darker states on the map have stronger positive correlations between fracking ALE and stroke mortality above 65. Figure 4 shows that the positive correlation between the ALE of fracking and stroke mortality (65+) in North Dakota, Ohio, Montana, Kansas, and Arkansas is stronger than in other states of fracking. Fracking has been active in North Dakota and Ohio [65]. Although fracking is generally active in Montana, Kansas, and Arkansas, fracking in their respective neighbors (e.g., North Dakota, Colorado, Texas) is always active. In contrast, the positive correlation in Virginia, Mississippi, and Oklahoma is weaker than in other states of fracking. Fracking is not active in the states of Virginia and Mississippi, according to the report by Environment America [4]. Additionally, Oklahoma, New Mexico, and Texas, where fracking is the most prevalent, have lower ALE coefficients for fracking than most states. The possible reason for this is



that Fracking ALE is positively correlated with some socioeconomic factors, such as family mean income and marital rate (see Figure A1), which are negatively correlated with stroke mortality over 65 years (see Figure A2).



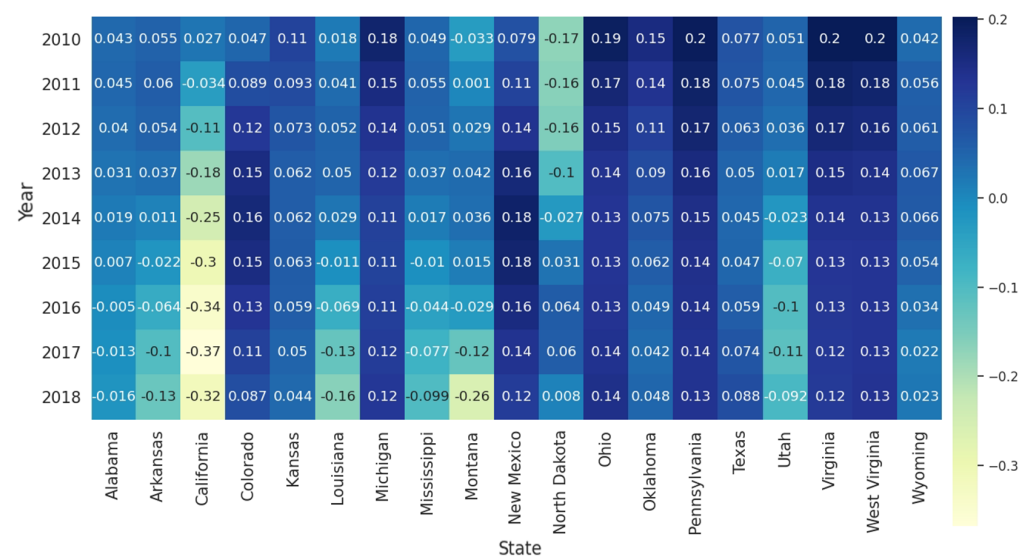
**Figure 4.** Spatial distribution of the average coefficients of fracking ALE for 19 states related to fracking.

### 3.3. Comparative Analysis on the Effect of Fracking on Gender-Based Stroke Mortality

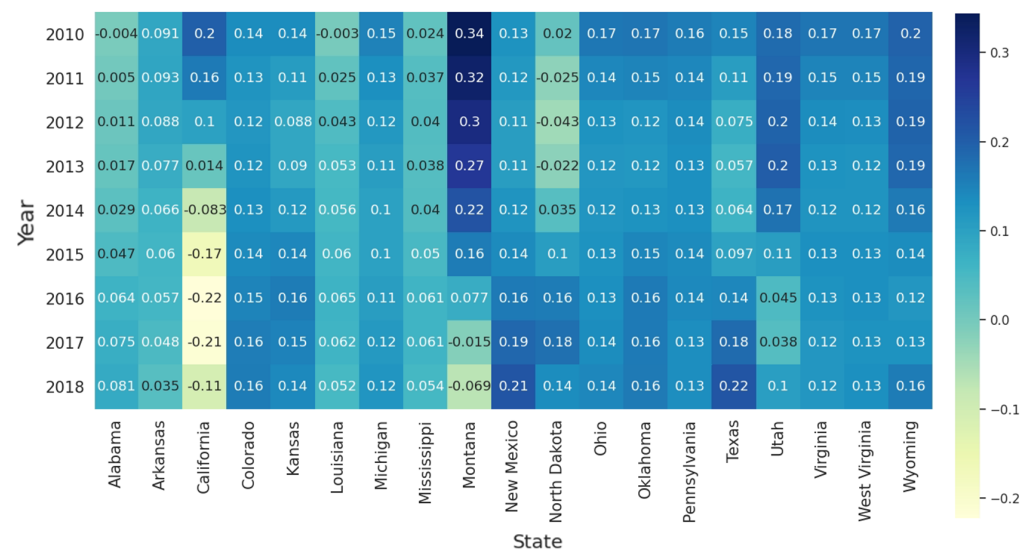
To explore spatio-temporal differences in the effect of hydraulic fracturing on gender-based stroke mortality, we performed spatio-temporal regressions on male and female stroke mortality separately using the selected explanatory variables. Based on the regression coefficients obtained for the two groups, the temporal and spatial characteristics of different sexes were compared and analyzed.

Figures 5 and 6 show, respectively, the regression coefficients of fracking ALE for different dependent variables (male stroke mortality and female stroke mortality). We found that the fracking ALE regression coefficients for males with stroke mortality and females with stroke mortality have similar temporal trends in Ohio, West Virginia, Virginia, and Pennsylvania. They slowly decreased from 2010 to 2018. The correlation coefficient between fracking ALE and stroke mortality (both men and women) in California decreased significantly year by year. The correlation coefficient between fracking ALE and stroke mortality (both male and female) in North Dakota increased significantly year by year. The correlation coefficients in the state of Colorado first increased and then decreased for male stroke mortality, but reversed for female stroke mortality. Additionally, the correlation coefficients between ALE from fracking and stroke mortality (both male and female) in California and some states (Arkansas, Louisiana, Mississippi, and Alabama) in the stroke belt are lower than those of other states from fracking.

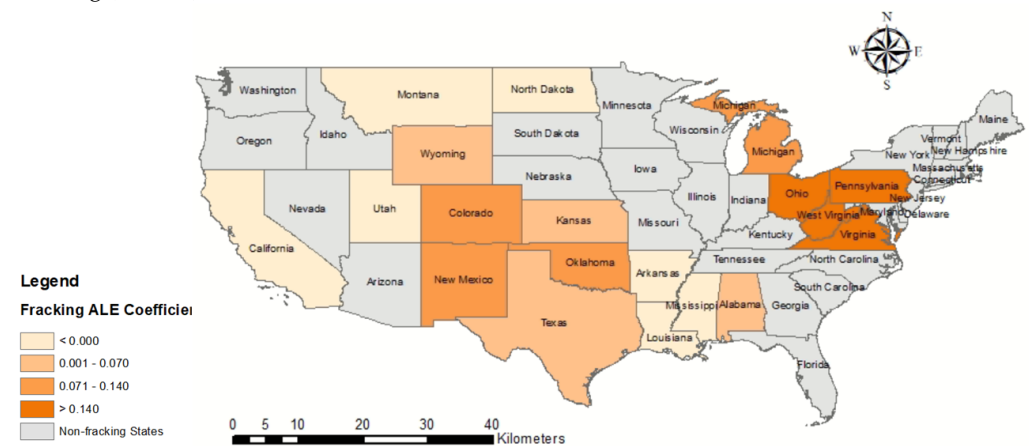
In addition to the temporal characteristics of gender-based stroke mortality, the fracking ALE regression coefficients are spatially differentiated (see Figure 7). The ALE coefficients of fracking for male and female on stroke mortality around Marcellus shale (including Pennsylvania, West Virginia, and Ohio) and New Mexico and Oklahoma are higher than those of other states of fracking. In contrast, the ALE coefficients for fracking on female stroke mortality are higher in Montana and Wyoming, but its ALE coefficients for male are lower.



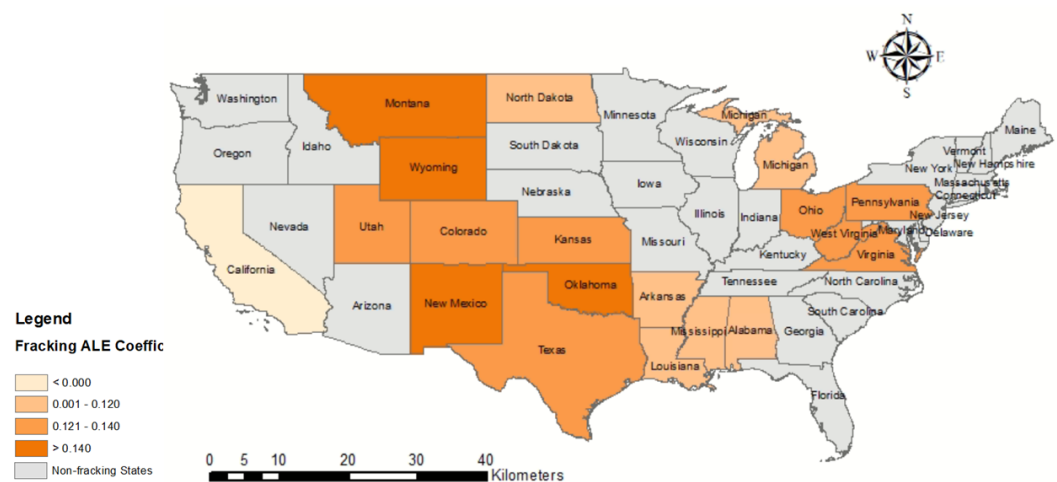
**Figure 5.** Temporal distribution of the average coefficients of the ALE of Fracking for 19 states related to fracking (Male).



**Figure 6.** Temporal distribution of the average coefficients of ALE of fracking for 19 states related to fracking (Female).



**Figure 7.** Cont.



**Figure 7.** Spatial distribution of the average coefficients of fracking ALE for 19 states related to fracking for Male (**upper**) and Female (**lower**).

#### 4. Discussion

##### 4.1. Does Fracking Cause a Higher Risk of Stroke?

Despite the potential health risks associated with fracking, there have been several quantitative studies on how fracking can affect public health on a local scale [40,66], characterizing the risk of fracking as the distance from the patient's residence to the nearest well. However, these methods are based on privacy-protected clinical data that contain large amounts of patient personal information. They are not applicable to this study, since patient personal information is not included in the publicly available dataset from CDC. This study uses ALE to quantitatively study the possible connection between fracking and stroke using publicly available data. It can be observed from the results of the GTWR analysis that fracking has a non-negligible effect on stroke mortality above 65 in most areas with fracking prevalent, as shown in Table 3 and Figure 4. According to USCB, there were 40.3 M US residents 65 years and older in the 2010 Census and 54.1 M in the population estimates of 1 July 2019, (<https://www.census.gov/topics/population/older-aging.html>). With the aging of the United States and the increase in fracking activities, how to keep fracking activities away from communities with a high proportion of adults 65 years or older is a question that policy makers should consider. However, its impact is relatively minor compared to other more dominant factors, such as cardiovascular disease and overdose.

##### 4.2. Spatiotemporal Differences in the Effect of Fracking on Age/Gender-Based Stroke Mortality

Deeper reasoning is needed to understand the geographic variations of cases of stroke mortality related to fracking based on the age and sex of the patients. As shown in Tables A1 and A2, the Fracking ALE variable for stroke mortality between 45–64 years is not considered statistically significant with a  $p$ -value  $> 0.1$ , but the Fracking ALE variable for stroke mortality over 65 years is considered statistically significant with a  $p$ -value  $< 0.01$ . For gender-based stroke mortality, the higher Fracking ALE coefficient for men than for women seems to suggest that men are at higher risk, at least in regions near Marcellus shale. Additionally, higher fracking ALE coefficients for women than for men suggest that women are at greater risk in some states, including Montana, Wyoming, Oklahoma, and New Mexico. Furthermore, the fracking ALE coefficients for both male and female stroke mortality are small in some states of the stroke belt (e.g., Arkansas, Louisiana, Mississippi, and Alabama), suggesting that fracking is not a major factor in stroke mortality compared to some major factors, such as cardiovascular disease.

##### 4.3. Air Pollutant Emissions from Fracking

We examined a possible mechanism for how fracking threatens public health. An existing report indicates that fracking produces environmental pollution, including hazardous

water pollutants and hazardous air pollutants (HAPs) [67]. To explore which HAPs are associated with fracking activities, we collected air pollutants data from the HAP monitoring station in Colorado (longitude:  $-108.053259$  and latitude:  $39.453654$ ). We have drawn a zone of interest with a radius of two kilometers around the HAP monitoring station and calculated the number of fracking wells. Four HAPs (Butadiene, Benzene, Formaldehyde, and Acetaldehyde) were monitored through this station, and we upsampled the four HAP monitoring data through linear interpolation to ensure that all four HAPs had the same time resolution. We then constructed a time series of fracking activities according to the start and end times of each fracking activity, and the overlapping fracking activities were aggregated. Finally, the four HAPs and the time series of the fracking activity were normalized, and then the fracking activities were analyzed for time series correlation with different HAPs using Pearson's correlation [68]. Pearson's  $r$ , which ranges from  $-1$  to  $1$ , was calculated to measure the degree of correlation between the two time series. When  $r > 0$ , a larger Pearson's  $r$  suggests a stronger positive correlation. Table 5 shows that the Pearson correlation ( $r$ ) of benzene was higher than the other three HAPs and the correlation is statistically significant ( $p$ -value  $< 0.01$ ). A study has shown the correlation between benzene exposure and the risk of cardiovascular disease due to the high level of trans, trans-muconic acid (t,t-MA) [69]. Although benzene exposure has also been found to be associated with high cholesterol [70], cardiovascular disease has a stronger correlation with stroke mortality at 65+ in our study than high cholesterol. The high density lipoprotein cholesterol was found to be more important for patients  $\leq 65$  years of age than older adults [71]. In addition, high cholesterol has been shown not to be associated with stroke mortality in some studies [33,72]. Therefore, the leakage of the chemical benzene due to fracking might contribute to cardiovascular disease thus stroke mortality but further study is needed and other processes such as water pollution might contribute as well.

**Table 5.** Pearson correlation significance test.

Air Pollutants	Pearson $r$	$p$ -Value
Butadiene	0.119	0.083
Formaldehyde	0.093	0.175
Acetaldehyde	0.049	0.474
Benzene	0.245	0.000 *

\* indicates  $p < 0.01$ .

#### 4.4. What Is the Implication of This Study on Health Policy-Making?

With the increase in fracking activity, the socioeconomic environment, such as the employment rate and family income, continues to increase. At the same time, more and more people and communities in areas with many hydraulically fractured wells report health problems, such as cancer and harm to the nervous, respiratory, and immune systems. Macroscopically, we found that the effects of fracking and stroke mortality were not significant for the 19 active states of fracking in the US. This is likely due to the following reasons: (1) Fracking areas in the US are mainly distributed in the Great Plains and Marcellus Shale, which are often located in mountainous or suburban areas with low population density, which may lead to limited spatial impact. (2) The fracking process generally only lasts 3–5 days, which leads to a limited impact on time. (3) The development of fracking will promote local socioeconomic status (e.g., employment rate and marital rate), which is negatively related to stroke mortality. This may cause the effect of fracking on stroke mortality to be insignificant in some states (e.g., Texas) where fracking is active.

This study analyzed the correlation between fracking and different hazardous air pollutants based on the public air pollutant dataset from the US Environmental Protection Agency (EPA). We found that the concentration of benzene in the air was related to the calculated sequence of fracking activity in the buffer zone within a radius of 2 km (Pearson  $r = 0.2452$  and the  $p$ -value  $< 0.01$ ). To minimize health risk, this research suggests that there should be no public facilities with a high population density within 2 km of fracking

activities. Furthermore, the high concentration of benzene in the air may be due to fracking, which caused groundwater pollution due to the extremely high volatility of benzene. It may be important to identify and investigate domestic water wells that are within two kilometers of a fracking well. A more systematic study of the impact of fracking on water contamination [28] remains for future research.

## 5. Conclusions

This article provided a systematic study on the spatiotemporal correlation between fracking and stroke mortality using the GTWR model. The temporal trend of positive correlation between fracking ALE and stroke mortality shows a varying pattern from state to state. The spatial distribution appears to demonstrate that there is a gender difference between the Great Plains and the Marcellus Shale. Our regression results also show that disease-related risk factors, including cardiovascular and overdose, have a more significant correlation with stroke mortality over 65 years of age than those related to fracking. Finally, there appears to be a significant temporal dependency between fracking and air pollutant emissions, especially for benzene. Future studies may focus on developing county-level GWR/GTWR models, although missing data is a critical challenge to resolve.

**Author Contributions:** Data curation, C.H.; Formal analysis, C.H.; Funding acquisition, S.W. and A.A.; Investigation, C.H. and S.W.; Methodology, C.H.; Project administration, S.W. and X.L.; Software, C.H.; Supervision, B.L. and X.L.; Writing—original draft, X.L.; Writing—review and editing, Z.Z., J.S. and X.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data generated or analyzed during this study are included in this article. Further inquiries can be addressed to the corresponding author.

**Acknowledgments:** This research is partially supported by NSF OIA-1920920 “R11 Track 2 FEC: Multi-Scale Integrative Approach to Digital Health: Collaborative Research and Education in Smart Health in West Virginia and Arkansas”.

**Conflicts of Interest:** None of the authors participating in this manuscript has conflict of interest to declare.

## Appendix A

The VIF values of the explanatory variables are shown in Table A1 (for stroke mortality older than 65 years) and Table A2 (for stroke mortality from 45 to 64 years). Figure A1 shows the correlation coefficient between two explanatory variables. The correlation coefficient ranges from  $-1$  to  $1$ . The closer the coefficient is to  $1$ , the stronger the positive correlation, the closer to  $-1$ , the stronger the negative correlation, and the closer to  $0$ , the smaller the correlation. To avoid multicollinearity, explanatory variables (e.g., hypertension and obesity) with VIF values greater than  $10$  have been removed. Table A3 shows the results of the Moran’s  $I$  test for explanatory variables. The VIF and Moran’s  $I$  values were used to identify explanatory variables involved in the regression analysis. Regression analysis of the GWR model based on the selected explanatory variables obtains regression coefficients for several significant explanatory variables. Some important explanatory variables are visualized in the space for comparative analysis with the ALE regression coefficients of fracking, as shown in Figure A2. Additionally, to investigate whether air pollutant emissions from Fracking affects stroke mortality, monitoring data for benzene concentrations in Colorado are shown in Figure A3.



**Table A1.** VIF values of explanatory variables of the OLS model for 65+ stroke mortality.

Variables	Fracking States			Non-Fracking States		
	Coef.	p-Value	VIF	Coef.	p-Value	VIF
<b>Related-disease risk factors</b>						
Diabetes	0.126	0.043	3.550	0.242	0.002	1.847
Cardiovascular	0.351	0.000	5.168	0.377	0.000	1.847
Overdose	0.531	0.000	6.309	0.239	0.012	3.698
<b>Behavior risk factors</b>						
Tobacco use	−0.079	0.255	5.193	0.076	0.478	2.959
High cholesterol	−0.179	0.002	3.284	−0.257	0.000	2.210
PAI	−0.032	0.572	3.698	−0.109	0.132	1.853
<b>Socioeconomic risk factors</b>						
Mean income	0.290	0.002	5.236	−0.191	0.009	2.961
Marital rate	−0.918	0.000	2.880	−0.204	0.042	2.078
Employment rate	−0.076	0.232	2.072	−0.309	0.000	1.598
<b>Fracking risk factors</b>						
Fracking ALE	0.113	0.000	1.511			

**Table A2.** VIF values of the explanatory variables of the OLS model for 45–64 stroke mortality.

Variables	Fracking States			Non-Fracking States		
	Coef.	p-Value	VIF	Coef.	p-Value	VIF
<b>Related-disease risk factors</b>						
Diabetes	0.304	0.000	3.550	0.2220	0.000	1.847
Cardiovascular	0.637	0.000	5.168	0.7286	0.000	1.847
Overdose	0.048	0.435	6.309	0.0376	0.513	3.698
<b>Behavior risk factors</b>						
Tobacco use	−0.016	0.784	5.193	0.1970	0.003	2.959
High cholesterol	−0.027	0.568	3.284	−0.0377	0.364	2.210
PAI	0.0626	0.187	3.698	0.1752	0.000	1.853
<b>Socioeconomic risk factors</b>						
Mean income	−0.040	0.612	5.236	−0.0682	0.121	2.961
Marital rate	0.093	0.360	2.880	0.3925	0.000	2.078
Employment rate	−0.236	0.000	2.072	−0.1005	0.037	1.598
<b>Fracking risk factors</b>						
Fracking ALE	0.0337	0.128	1.511			

**Table A3.** Moran's I test result for explanatory variables.

Variables	Fracking States			Non-Fracking States		
	Moran's I	Z-Score	p-Value	Moran's I	Z-Score	p-Value
<b>Related-disease risk factors</b>						
Diabetes	0.799	24.055	0.000	0.659	20.274	0.000
Cardiovascular	0.955	28.695	0.000	0.924	28.348	0.000
Overdose	0.725	21.836	0.000	0.666	20.501	0.000
<b>Behavior risk factors</b>						
Tobacco use	0.854	25.731	0.000	0.589	18.140	0.000
High cholesterol	0.553	16.679	0.000	0.321	9.944	0.000
PAI	0.925	27.811	0.000	0.759	23.426	0.000
<b>Socioeconomic risk factors</b>						
Mean income	0.734	22.111	0.000	0.778	23.928	0.000
Marital rate	0.221	6.761	0.000	0.356	11.219	0.000
Employment rate	0.944	28.664	0.000	0.918	28.173	0.000
<b>Fracking risk factor</b>						
Fracking ALE	0.271	8.955	0.000			

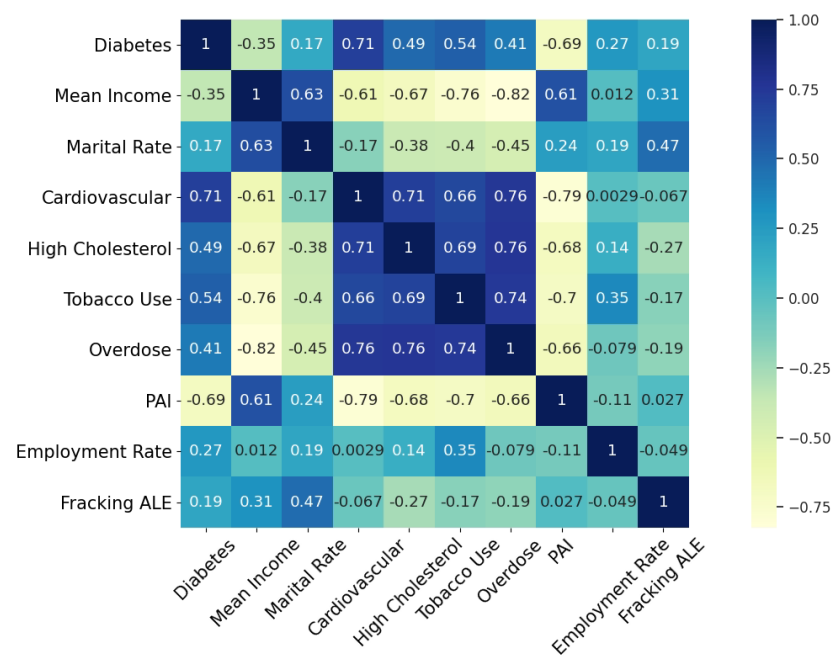


Figure A1. VIF between explanatory variables.

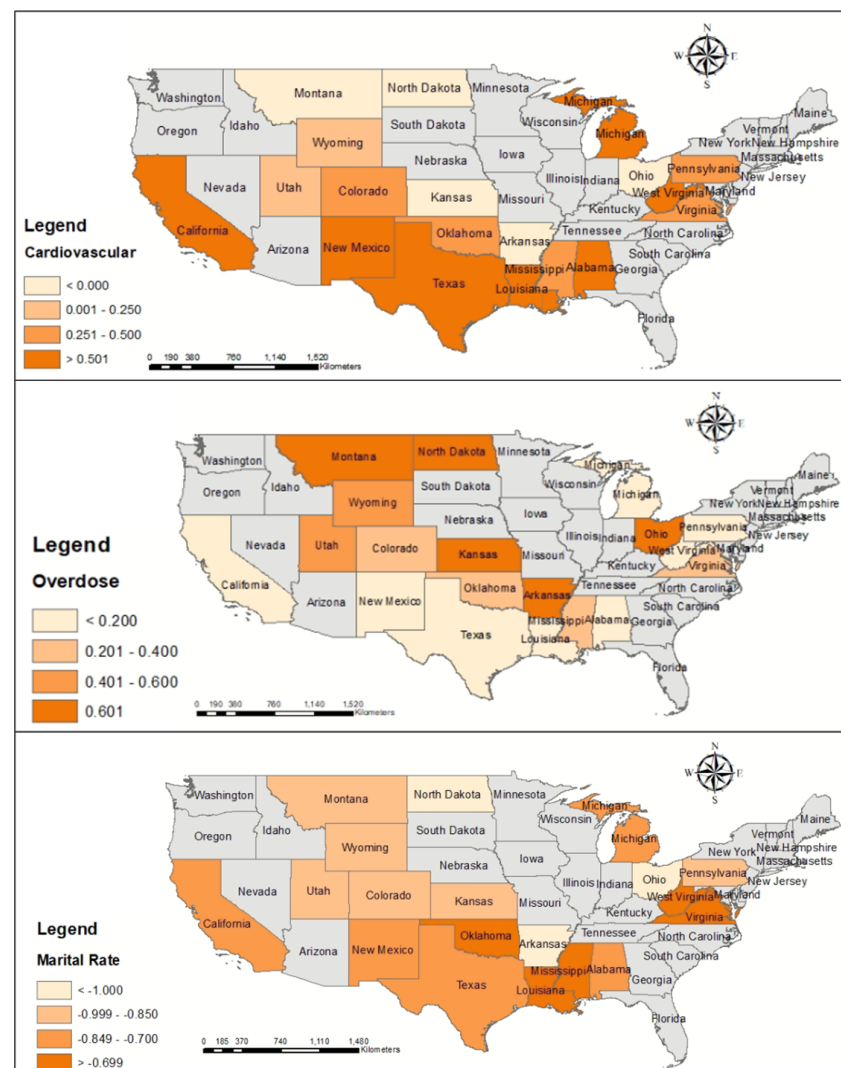


Figure A2. Spatial distribution of the significant coefficients for 19 states related to fracking.

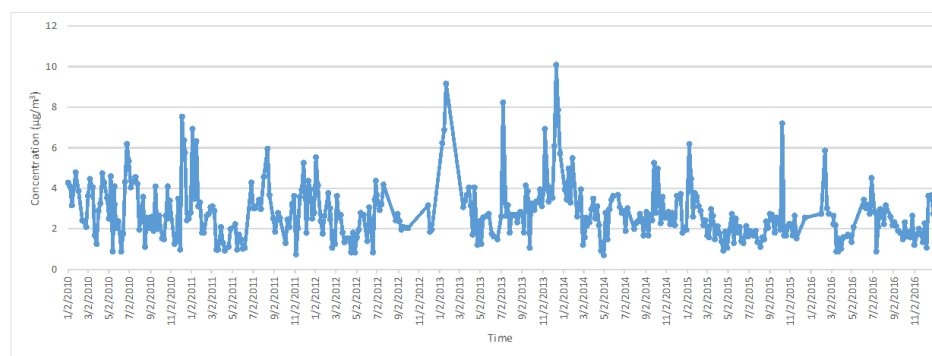


Figure A3. Benzene concentration monitoring data.

## Reference

- Gandossi, L.; Von Estorff, U. An overview of hydraulic fracturing and other formation stimulation technologies for shale gas production. *Eur. Commisison Jt. Res. Cent. Tech. Rep.* **2013**, 26347.
- Montgomery, C.T.; Smith, M.B. Hydraulic fracturing: History of an enduring technology. *J. Pet. Technol.* **2010**, *62*, 26–40. [\[CrossRef\]](#)
- Kuuskraa, V.; Stevens, S.H.; Moodhe, K.D. *Technically Recoverable Shale Oil and Shale Gas Resources: An Assessment of 137 Shale Formations in 41 Countries Outside the United States*; US Energy Information Administration, US Department of Energy: Washington, DC, USA, 2013.
- U.S. Energy Information Administration. *Hydraulic Fracturing Accounts for About Half of Current U.S. Crude Oil Production*; U.S. Energy Information Administration: Washington, DC, USA, 2016.
- Finkel, M.L.; Hays, J. The implications of unconventional drilling for natural gas: A global public health concern. *Public Health* **2013**, *127*, 889–893. [\[CrossRef\]](#) [\[PubMed\]](#)
- Meng, Q. Characterizing and modeling environmental emergency of unconventional oil and gas spills in the USA: Life-year versus spill factors. *J. Clean. Prod.* **2019**, *237*, 117794. [\[CrossRef\]](#)
- Kovats, S.; Depledge, M.; Haines, A.; Fleming, L.E.; Wilkinson, P.; Shonkoff, S.B.; Scovronick, N. The health implications of fracking. *Lancet* **2014**, *383*, 757. [\[CrossRef\]](#)
- Black, K.J.; Boslett, A.J.; Hill, E.L.; Ma, L.; McCoy, S.J. Economic, Environmental, and Health Impacts of the Fracking Boom. *Annu. Rev. Resour. Econ.* **2021**, *13*, 311–334. [\[CrossRef\]](#)
- Soyer, M.; Kaminski, K.; Ziyanak, S. Socio-Psychological Impacts of Hydraulic Fracturing on Community Health and Well-Being. *Int. J. Environ. Res. Public Health* **2020**, *17*, 1186. [\[CrossRef\]](#)
- Caron-Beaudoin, É.; Whyte, K.P.; Bouchard, M.F.; Chevrier, J.; Haddad, S.; Copes, R.; Fröhlich, K.L.; Dokkie, D.; Juul, S.; Bouchard, M.; et al. Volatile organic compounds (VOCs) in indoor air and tap water samples in residences of pregnant women living in an area of unconventional natural gas operations: Findings from the EXPERIVA study. *Sci. Total. Environ.* **2022**, *805*, 150242. [\[CrossRef\]](#)
- Ma, L.; Hurtado, A.; Eguilior, S.; Borrajo, J.F.L. Exposure risk assessment to organic compounds based on their concentrations in return water from shale gas developments. *Sci. Total. Environ.* **2022**, *822*, 153586. [\[CrossRef\]](#)
- Meng, Q. The impacts of fracking on the environment: A total environmental study paradigm. *Sci. Total. Environ.* **2017**, *580*, 953–957. [\[CrossRef\]](#)
- McKenzie, L.M.; Witter, R.Z.; Newman, L.S.; Adgate, J.L. Human health risk assessment of air emissions from development of unconventional natural gas resources. *Sci. Total. Environ.* **2012**, *424*, 79–87. [\[CrossRef\]](#)
- Vengosh, A.; Jackson, R.B.; Warner, N.; Darrah, T.H.; Kondash, A. A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States. *Environ. Sci. Technol.* **2014**, *48*, 8334–8348. [\[CrossRef\]](#)
- Deziel, N.C.; Brokovich, E.; Grotto, I.; Clark, C.J.; Barnett-Itzhaki, Z.; Broday, D.; Agay-Shay, K. Unconventional oil and gas development and health outcomes: A scoping review of the epidemiological research. *Environ. Res.* **2020**, *182*, 109124. [\[CrossRef\]](#)
- Tao, Z.; Liu, C.; He, Q.; Chang, H.; Ma, J. Detection and treatment of organic matters in hydraulic fracturing wastewater from shale gas extraction: A critical review. *Sci. Total. Environ.* **2022**, *824*, 153887. [\[CrossRef\]](#)
- Meng, Q. Rethink potential risks of toxic emissions from natural gas and oil mining. *Environ. Pollut.* **2018**, *240*, 848–857. [\[CrossRef\]](#)
- Casey, J.A.; Savitz, D.A.; Rasmussen, S.G.; Ogburn, E.L.; Pollak, J.; Mercer, D.G.; Schwartz, B.S. Unconventional natural gas development and birth outcomes in Pennsylvania, USA. *Epidemiology* **2016**, *27*, 163. [\[CrossRef\]](#)
- 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*, 1334–1343. [\[CrossRef\]](#)

20. 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.* **2017**, *125*, 189–197. [\[CrossRef\]](#)
21. 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. [\[CrossRef\]](#)
22. Marlin-Tackie, F.A.; Smith, J.M. Key characteristics influencing risk perceptions of unconventional energy development. *J. Clean. Prod.* **2020**, *251*, 119644. [\[CrossRef\]](#)
23. McDermott-Levy, R.; Kaktins, N.; Sattler, B. Fracking, the environment, and health. *AJN Am. J. Nurs.* **2013**, *113*, 45–51. [\[CrossRef\]](#)
24. Wong, C.M.; Vichit-Vadakan, N.; Kan, H.; Qian, Z. Public Health and Air Pollution in Asia (PAPA): a multicity study of short-term effects of air pollution on mortality. *Environ. Health Perspect.* **2008**, *116*, 1195–1202. [\[CrossRef\]](#)
25. Kelly, F.J.; Fussell, J.C. Air pollution and public health: emerging hazards and improved understanding of risk. *Environ. Geochem. Health* **2015**, *37*, 631–649. [\[CrossRef\]](#)
26. Hu, C.B.; Zhang, F.; Gong, F.Y.; Ratti, C.; Li, X. Classification and mapping of urban canyon geometry using Google Street View images and deep multitask learning. *Build. Environ.* **2020**, *167*, 106424. [\[CrossRef\]](#)
27. Turner, M.C.; Andersen, Z.J.; Baccarelli, A.; Diver, W.R.; Gapstur, S.M.; Pope, C.A., III; Prada, D.; Samet, J.; Thurston, G.; Cohen, A. Outdoor air pollution and cancer: An overview of the current evidence and public health recommendations. *CA Cancer J. Clin.* **2020**, *70*, 460–479. [\[CrossRef\]](#)
28. Jackson, R.B.; Vengosh, A.; Carey, J.W.; Davies, R.J.; Darrah, T.H.; O'sullivan, F.; Pétron, G. The environmental costs and benefits of fracking. *Annu. Rev. Environ. Resour.* **2014**, *39*, 327–362. [\[CrossRef\]](#)
29. Stacy, S.L. A review of the human health impacts of unconventional natural gas development. *Curr. Epidemiol. Rep.* **2017**, *4*, 38–45. [\[CrossRef\]](#)
30. Towfighi, A.; Saver, J.L. Stroke declines from third to fourth leading cause of death in the United States: Historical perspective and challenges ahead. *Stroke* **2011**, *42*, 2351–2355. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Lanska, D.J.; Kuller, L.H. The geography of stroke mortality in the United States and the concept of a stroke belt. *Stroke* **1995**, *26*, 1145–1149. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Liao, Y.; Greenlund, K.J.; Croft, J.B.; Keenan, N.L.; Giles, W.H. Factors explaining excess stroke prevalence in the US Stroke Belt. *Stroke* **2009**, *40*, 3336–3341. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Feigin, V.L.; Roth, G.A.; Naghavi, M.; Parmar, P.; Krishnamurthi, R.; Chugh, S.; Mensah, G.A.; Norrving, B.; Shieue, I.; Ng, M.; et al. Global burden of stroke and risk factors in 188 countries, during 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. *Lancet Neurol.* **2016**, *15*, 913–924. [\[CrossRef\]](#)
34. Howard, G.; Howard, V.J. Twenty years of progress toward understanding the stroke belt. *Stroke* **2020**, *51*, 742–750. [\[CrossRef\]](#)
35. Srebotnjak, T.; Rotkin-Ellman, M. Fracking fumes: Air pollution from hydraulic fracturing threatens public health and communities. *Nat. Resour. Def. Coun.* **2014**. Available online: [https://legacy-assets.eenews.net/open\\_files/assets/2014/12/16/document\\_cw\\_01.pdf](https://legacy-assets.eenews.net/open_files/assets/2014/12/16/document_cw_01.pdf) (accessed on 25 August 2022).
36. Law, A.; Duff, D.; Saunders, P.; Middleton, J.; McCoy, D. Medical organisations must divest from fossil fuels. *BMJ* **2018**, *363*, k5163. [\[CrossRef\]](#)
37. Krajnak, K.; Kan, H.; Russ, K.A.; McKinney, W.; Waugh, S.; Zheng, W.; Kashon, M.L.; Johnson, C.; Cumpston, J.; Fedan, J.S. Biological effects of inhaled hydraulic fracturing sand dust. VI. Cardiovascular effects. *Toxicol. Appl. Pharmacol.* **2020**, *406*, 115242. [\[CrossRef\]](#)
38. Oikonomou, P.D.; Kallenberger, J.A.; Waskom, R.M.; Boone, K.K.; Plombon, E.N.; Ryan, J.N. Water acquisition and use during unconventional oil and gas development and the existing data challenges: Weld and Garfield counties, CO. *J. Environ. Manag.* **2016**, *181*, 36–47. [\[CrossRef\]](#)
39. Purvis, R.M.; Lewis, A.C.; Hopkins, J.R.; Wilde, S.E.; Dunmore, R.E.; Allen, G.; Pitt, J.; Ward, R.S. Effects of ‘pre-fracking’ operations on ambient air quality at a shale gas exploration site in rural North Yorkshire, England. *Sci. Total. Environ.* **2019**, *673*, 445–454. [\[CrossRef\]](#)
40. Currie, J.; Greenstone, M.; Meckel, K. Hydraulic fracturing and infant health: New evidence from Pennsylvania. *Sci. Adv.* **2017**, *3*, e1603021. [\[CrossRef\]](#)
41. Meng, Q. Spatial analysis of environment and population at risk of natural gas fracking in the state of Pennsylvania, USA. *Sci. Total. Environ.* **2015**, *515*, 198–206. [\[CrossRef\]](#)
42. Meng, Q.; Ashby, S. Distance: A critical aspect for environmental impact assessment of hydraulic fracking. *Extr. Ind. Soc.* **2014**, *1*, 124–126. [\[CrossRef\]](#)
43. Brunsdon, C.; Fotheringham, A.S.; Charlton, M.E. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geogr. Anal.* **1996**, *28*, 281–298. [\[CrossRef\]](#)
44. Huang, B.; Wu, B.; Barry, M. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 383–401. [\[CrossRef\]](#)
45. Feigin, V.L.; Forouzanfar, M.H.; Krishnamurthi, R.; Mensah, G.A.; Connor, M.; Bennett, D.A.; Moran, A.E.; Sacco, R.L.; Anderson, L.; Truelsen, T.; et al. Global and regional burden of stroke during 1990–2010: Findings from the Global Burden of Disease Study 2010. *Lancet* **2014**, *383*, 245–255. [\[CrossRef\]](#)

46. Patterson, L.A.; Kohnschnik, K.E.; Wiseman, H.; Fargione, J.; Maloney, K.O.; Kiesecker, J.; Nicot, J.P.; Baruch-Mordo, S.; Entrekin, S.; Trainor, A.; et al. Unconventional oil and gas spills: Risks, mitigation priorities, and state reporting requirements. *Environ. Sci. Technol.* **2017**, *51*, 2563–2573.
47. Korfmacher, K.S.; Jones, W.A.; Malone, S.L.; Vinci, L.F. Public health and high volume hydraulic fracturing. *New Solut. J. Environ. Occup. Health Policy* **2013**, *23*, 13–31.
48. Colborn, T.; Kwiatkowski, C.; Schultz, K.; Bachran, M. Natural gas operations from a public health perspective. *Hum. Ecol. Risk Assess. Int. J.* **2011**, *17*, 1039–1056.
49. Bamberger, M.; Oswald, R.E. Impacts of gas drilling on human and animal health. *New Solut. J. Environ. Occup. Health Policy* **2012**, *22*, 51–77. [\[CrossRef\]](#)
50. Bistarelli, S.; Fioravanti, F.; Peretti, P. Defense trees for economic evaluation of security investments. In Proceedings of the First International Conference on Availability, Reliability and Security (ARES'06), Vienna, Austria, 20–22 April 2006; p. 8.
51. Karabacak, B.; Sogukpinar, I. ISRAM: Information security risk analysis method. *Comput. Secur.* **2005**, *24*, 147–159.
52. Akinwande, M.O.; Dikko, H.G.; Samson, A. Variance inflation factor: As a condition for the inclusion of suppressor variable (s) in regression analysis. *Open J. Stat.* **2015**, *5*, 754. [\[CrossRef\]](#)
53. Kutner, M.H.; Nachtsheim, C.J.; Neter, J.; Wasserman, W. *Applied Linear Regression Models*; McGraw-Hill/Irwin: New York, NY, USA, 2004; Volume 4.
54. Moran, P.A. Notes on continuous stochastic phenomena. *Biometrika* **1950**, *37*, 17–23. [\[CrossRef\]](#)
55. Shekhar, S.; Jiang, Z.; Ali, R.Y.; Eftelioglu, E.; Tang, X.; Gunturi, V.M.; Zhou, X. Spatiotemporal data mining: A computational perspective. *ISPRS Int. J. Geo-Inf.* **2015**, *4*, 2306–2338. [\[CrossRef\]](#)
56. Wang, W.; Hu, C.; Chen, N.; Xiao, C.; Wang, C.; Chen, Z. Spatio-temporal enabled urban decision-making process modeling and visualization under the cyber-physical environment. *Sci. China Inf. Sci.* **2015**, *58*, 1–17. [\[CrossRef\]](#)
57. Xiao, C.; Chen, N.; Gong, J.; Wang, W.; Hu, C.; Chen, Z. Event-driven distributed information resource-focusing service for emergency response in smart city with cyber-physical infrastructures. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 251.
58. Chuanbo, H.; Lan, Y.; Hui, L. A Process-oriented Risk Monitoring and Assessment Modeling Approach in Urban Public Safety. *Acta Geod. Cartogr. Sin.* **2018**, *47*, 1062.
59. Wang, W.; Hu, C.; Chen, N.; Xiao, C.; Jia, S. Spatio-Temporal Risk Assessment Process Modeling for Urban Hazard Events in Sensor Web Environment. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 203. [\[CrossRef\]](#)
60. Cavanaugh, J.E. Unifying the derivations for the Akaike and corrected Akaike information criteria. *Stat. Probab. Lett.* **1997**, *33*, 201–208. [\[CrossRef\]](#)
61. Dismuke, C.; Lindrooth, R. Ordinary least squares. *Methods Des. Outcomes Res.* **2006**, *93*, 93–104.
62. Arboix, A. Cardiovascular risk factors for acute stroke: Risk profiles in the different subtypes of ischemic stroke. *World J. Clin. Cases WJCC* **2015**, *3*, 418.
63. Avan, A.; Digaleh, H.; Di Napoli, M.; Stranges, S.; Behrouz, R.; Shojaeianbabaei, G.; Amiri, A.; Tabrizi, R.; Mokhber, N.; Spence, J.D.; et al. Socioeconomic status and stroke incidence, prevalence, mortality, and worldwide burden: An ecological analysis from the Global Burden of Disease Study 2017. *BMC Med.* **2019**, *17*, 191. [\[CrossRef\]](#)
64. Addo, J.; Ayerbe, L.; Mohan, K.M.; Crichton, S.; Sheldenkar, A.; Chen, R.; Wolfe, C.D.; McKevitt, C. Socioeconomic status and stroke: An updated review. *Stroke* **2012**, *43*, 1186–1191.
65. Center, E.A.R.P. *Fracking by the Numbers The Damage to Our Water, Land and Climate from a Decade of Dirty Drilling*; Environment America Research and Policy Center: Washington, DC, USA, 2016.
66. Clark, C.J.; Xiong, B.; Soriano, M.A., Jr.; Gutches, K.; Siegel, H.G.; Ryan, E.C.; Johnson, N.P.; Cassell, K.; Elliott, E.G.; Li, Y.; et al. Assessing unconventional oil and gas exposure in the Appalachian basin: Comparison of exposure surrogates and residential drinking water measurements. *Environ. Sci. Technol.* **2022**, *56*, 1091–1103. [\[CrossRef\]](#)
67. Windham, G.C.; Zhang, L.; Gunier, R.; Croen, L.A.; Grether, J.K. Autism spectrum disorders in relation to distribution of hazardous air pollutants in the San Francisco Bay area. *Environ. Health Perspect.* **2006**, *114*, 1438–1444. [\[CrossRef\]](#)
68. Benesty, J.; Chen, J.; Huang, Y.; Cohen, I. Pearson correlation coefficient. In *Noise Reduction in Speech Processing*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 1–4.
69. Abplanalp, W.; DeJarnett, N.; Riggs, D.W.; Conklin, D.J.; McCracken, J.P.; Srivastava, S.; Xie, Z.; Rai, S.; Bhatnagar, A.; O'Toole, T.E. Benzene exposure is associated with cardiovascular disease risk. *PLoS ONE* **2017**, *12*, e0183602. [\[CrossRef\]](#)
70. Alshareef, A.; Ibrahim, M. Neurological and Biological Toxicity of Subchronic Exposure to Inhaled Benzene in Male Rats. *J. Biochem. Technol.* **2020**, *11*, 52–59.
71. Tirschwell, D.; Smith, N.; Heckbert, S.; Lemaitre, R.; Longstreth, W.; Psaty, B. Association of cholesterol with stroke risk varies in stroke subtypes and patient subgroups. *Neurology* **2004**, *63*, 1868–1875. [\[CrossRef\]](#)
72. Bowman, T.S.; Sesso, H.D.; Ma, J.; Kurth, T.; Kase, C.S.; Stampfer, M.J.; Gaziano, J.M. Cholesterol and the risk of ischemic stroke. *Stroke* **2003**, *34*, 2930–2934. [\[CrossRef\]](#)