Health Guard: Health and Preventive Medicine Research

PRIME OPEN ACCESS

Volume 1, Issue 1

Research Article

Date of Submission: 25 April, 2025
Date of Acceptance: 22 July 2025
Date of Publication: 08 August 2025

The Role of Socioeconomic, Sociodemographic, Weather, and Health Factors in Influencing Covid-19 Cases and Deaths in Oklahoma

Amune Ofuje Blessing* and Jonathan C Comer

Department of Geography, Oklahoma State University, USA

***Corresponding Author:**
Amune Ofuje Blessing, Department of Geography, Oklahoma State University, USA.

Citation: Blessing, A. O., Comer, J. C. (2025). The Role of Socioeconomic, Sociodemographic, Weather, and Health Factors in Influencing Covid-19 Cases and Deaths in Oklahoma. *HealthGuard J Health Prev Med, 1*(1), 01-19.

Abstract

Current research on COVID-19 has highlighted the importance of analyzing the relationship between various risks factors that may contribute to the proliferation of COVID-19 cases and deaths in some areas than others. The aim of this study is to identify the various risk factors that may have contributed to the geographical distribution COVID-19 cases and deaths in Oklahoma. The results of the study could help provide targeted interventions and policies that could address the needs of communities and help provide more effective control of the virus.

Two major methods were used to analyze the data, MGWR and ZINB model. The MGWR was used to analyze the COVID-19 cases and the ZINB was used to analyze the COVID-19 deaths. The result of the study suggests that "those with health conditions, humidity and no solar radiation, those with poor economic condition, and non-English speaking young Hispanics, had significant association with COVID-19 case in Oklahoma but in varied location. For COVID-19 deaths, only those with health conditions, humidity and no solar radiation, and those with poor economic condition, had significant association with COVID-19 deaths. The results of the study can be useful to policy makers and healthcare providers in understanding regional differences when making policies as well as understanding the importance of applying local policies as a better strategy to addressing health related problems.

Introduction

The recent COVID-19 pandemic has brought to the forefront the importance of understanding the spatial and spatiotemporal patterns of infectious diseases. There has been a growing interest in examining the role of socioeconomic, sociodemographic, weather, and health factors in explaining differences and changes in the spatial and spatiotemporal patterns of COVID-19. Such analysis is crucial for the development of effective policies to prevent and control the spread of the virus, and future similar occurrences especially in areas with higher cases of COVID-19. The Geography of COVID-19 in the United States In the United States, studies have found a disproportional distribution of COVID-19 cases, meaning that it does not affect everyone equally. The US COVID-19 cases expose the inequalities in the distribution of resources and health care systems in the United States. Confirming the social inequalities in the US system, studies have found that Black populations and other minority groups have been affected by the COVID-19 pandemic more than their counterparts [1].

In addition, the US deaths due to COVID-19 have been reported to be disproportionally higher among Blacks. For example, in Milwaukee, Wisconsin, three-quarters of COVID-19 related deaths were reported among Blacks [1]. This makes the US a special place to study with regards to the transmissibility of COVID-19, given that the current pandemic issues would worsen the existing racial, health and socio-economic inequalities problem. The country has the highest COVID-19 cases and related deaths worldwide, with about 103,268,408 confirmed cases and 1,115,637 deaths reported

as of February 23, 2023, before they stopped reporting data publicly [2].

However, between February 17 and April 6, 2024, the country had a reported total hospitalization of 6,917,410 and total deaths of 1,188,278 [3]. The disproportional differences in the distribution of COVID-19 cases and deaths in the US have been attributed to the structural factors preventing minorities and socially deprived groups practicing social distancing and taking general preventive measures. The Black community makes up "essential workers" in such businesses as grocery stores, public transit, health care, the country and custodial staff that puts them at a higher risk of contracting the virus and transmitting it easily to other people based on their social lifestyle and poor living conditions [3].

The disparity in health care systems in the US, especially with regards to health insurance, leaves many in jeopardy, particularly those with underlying health issues. In addition, weather has been discovered to have both a direct and indirect influence on health disparities in the US. For example, high mortality has been discovered to occur on extreme high temperature in urban areas, with disproportional risk among specific population groups . People with low socioeconomic status are unable to afford protective measures, such as air conditioning, leading to disproportional risks among the population. These disproportional conditions in the US outlined above make it a unique place to study the current pandemic situation, especially because the US has been identified as consistently having the highest number of COVID-19 cases and deaths globally since April 2020 [4].

To this end, current research on COVID-19 has highlighted the importance of analyzing the relationship between various risk factors such as age, temperature, diabetes, neighborhood deprivation, and others, in understanding the spatiotemporal patterns of the virus. These risk factors are known to play a key role in the transmission and spread of infectious diseases and could be useful in developing effective policies for controlling the virus. The risk of severe COVID-19 illness and death increases with age, particularly for those 65 years and older [5,6].

Underlying medical conditions such as diabetes, obesity, cardiovascular disease, chronic respiratory disease, and cancer are at higher risk of severe COVID-19 illness [5]. Studies have shown that people from racial and ethnic minority groups, particularly Black, Hispanic, and Indigenous populations, are disproportionately affected by COVID-19, both in terms of infection rates and mortality [7-11]. This disparity is partly due to underlying health conditions that are more common among these groups, such as diabetes, obesity, and hypertension. People with disabilities are also at higher risk of getting COVID-19 and experiencing severe complications [12].

This is due to a combination of factors including underlying health conditions, living in shared group settings, and barriers in accessing healthcare. According to the Centers for Disease Control and Prevention (CDC), individuals with weakened immune systems, such as those with HIV, cancer, or organ transplant recipients, may be at higher risk of severe COVID-19 illness [5]. In addition, men may be at higher risk of severe COVID-19 illness and death than women [13-15].

While there are established risk factors for COVID-19, it is also important to understand how these risk factors manifest in specific communities or geographic regions. This can help inform targeted interventions and policies that address the unique needs of those communities and can lead to more effective outcomes in reducing the spread and impact of COVID-19 [16,17]. For example, in areas with high population density, policies such as mask mandates and physical distancing may be particularly effective in reducing transmission. In areas with high rates of underlying health conditions, policies focused on improving access to healthcare and promoting healthy lifestyles may be effective in reducing the severity of COVID-19 illnesses.

Understanding the specific risk factors and needs of communities can also help address issues related to vaccine hesitancy and mistrust in public health interventions. By working with trusted community leaders and using culturally appropriate messaging, public health officials can help increase vaccine uptake and reduce the impact of COVID-19 in specific communities. Therefore, the main objective of this study is to determine the various socioeconomic, sociodemographic, weather, and health factors that are significantly associated with COVID-19 cases and deaths in Oklahoma.

To achieve this objective, the study uses spatial analysis techniques to analyze data on various established social risk factors, such as poverty levels, temperature, humidity, diabetes, and disability to determine their associations with COVID-19 cases and deaths. By identifying these factors, the study aims to provide insights into the drivers of COVID-19 transmission in Oklahoma and to inform targeted public health interventions to reduce the burden of the disease in the state.

Specifically, the study addresses two main research questions
- What social risk factors are associated with COVID -19 cases in Oklahoma?
- What social risk factors are associated with COVID -19 deaths in Oklahoma?

Literature Review

Introduction

For transmitted infections, viruses transmitted from human to human such as COVID-19, human contact patterns, and

population demographics are useful transmissibility measures to better understand the virus's spatial spread. In addition, underlying diseases such as diabetes, asthma, and other respiratory tract diseases may influence the transmission of COVID-19, and the large number of COVID-19 related deaths recorded. Hence, socio-environmental, socio-economic, health, weather and socio-demographic factors have been found to be a very useful determinant of the spatial spread and pattern of COVID-19 [18,19].

Socio-Economic Factors and COVID-19

The COVID-19 pandemic has highlighted the complex relationship between socio-economic factors and health outcomes. Research has shown that the virus has disproportionately affected marginalized communities, including low-income individuals, people of color, low educational status, and those living in overcrowded housing conditions or working in frontline jobs that are more prone to exposure to the virus [20-22].

One key factor contributing to the disparities in COVID-19 outcomes is the unequal access to healthcare resources [23]. Many low-income and minority communities lack access to quality healthcare services, making it difficult for them to receive timely and adequate care for COVID-19 [8]. Additionally, these communities often have higher rates of pre-existing health conditions, such as obesity, diabetes, and cardiovascular disease, which can increase the risk of severe COVID-19 outcomes [24].

Another factor is the nature of work and living conditions in these communities. Many low-wage workers, such as those in the service and hospitality industries, have been deemed essential during the pandemic and are more likely to be exposed to the virus. Moreover, many of these workers do not have access to paid sick leave, which can make it difficult for them to self-isolate and prevent the spread of the virus. Similarly, people living in overcrowded or substandard housing conditions may have difficulty practicing physical distancing or self-isolating when necessary [25,26].

Finally, the pandemic has also highlighted the role of systemic racism and discrimination in exacerbating health disparities. For example, racial and ethnic minorities in the United States are more likely to live in polluted areas with poor air quality, which can contribute to respiratory problems and increase the risk of severe COVID-19 outcomes. Additionally, many of these communities have experienced historical and ongoing discrimination, leading to increased stress, trauma, and other mental health challenges that can impact their ability to cope with the pandemic[1,23,27].

In conclusion, socio-economic factors play a significant role in determining COVID-19 outcomes. Addressing these disparities requires a comprehensive approach that addresses the underlying social determinants of health, such as access to quality healthcare, safe and affordable housing, and fair and equitable employment opportunities.

Health Factors and COVID-19

The COVID-19 pandemic has had a significant impact on public health, with a wide range of health factors influencing the likelihood of individuals contracting the virus and experiencing severe outcomes. Research has shown that several pre-existing health conditions increase the risk of severe COVID-19 outcomes, including obesity, diabetes, cardiovascular disease, chronic lung disease, and weakened immune systems. One factor that can impact COVID-19 outcomes is obesity [28-31].

Obesity has been associated with increased inflammation and a higher risk of chronic diseases such as diabetes and cardiovascular disease, both of which are linked to worse COVID-19 outcomes. Research has also shown that obesity may weaken the immune system's response to the virus, making it more difficult to fight off infection. Diabetes is another health factor that can impact COVID-19 outcomes. Individuals with diabetes have higher levels of inflammation and a weakened immune system, which can increase the risk of severe COVID-19 outcomes [32-34].

Additionally, uncontrolled blood sugar levels can make it more difficult for individuals to recover from the virus. Cardiovascular disease, including hypertension and heart failure, has also been linked to worse COVID-19 outcomes [35]. Individuals with these conditions have a higher risk of developing severe respiratory complications and requiring hospitalization. Chronic lung disease, such as chronic obstructive pulmonary disease (COPD) and asthma can also increase the risk of severe COVID-19 outcomes. These conditions can make it more difficult for individuals to breathe, putting them at higher risk of respiratory failure.

Finally, weakened immune systems can also contribute to worse COVID-19 outcomes. Individuals with conditions that weaken the immune system, such as cancer or HIV, may have a more challenging time fighting off the virus [5,30]. In conclusion, several pre-existing health conditions can increase the risk of severe COVID-19 outcomes. Addressing these conditions through preventative measures and appropriate medical treatment can help reduce the risk of severe COVID-19 outcomes. Additionally, public health efforts to promote healthy behaviors and reduce the prevalence of chronic diseases can help mitigate the impact of COVID-19 on public health.

Weather Factors and COVID-19

Several respiratory illnesses caused by viruses, such as influenza and other types of coronaviruses, exhibit a seasonal pattern that is partly due to the impact of weather conditions on virus survival, seasonal immunity, and population

mixing. Epidemiological studies have indicated that the risk of transmission of SARS-CoV-2, the virus responsible for COVID-19, is greater at lower ambient temperatures and lower humidity levels [26].

Weather and climate conditions have been found to have a significant influence on the transmission and prevalence of infectious diseases. The COVID-19 pandemic is no exception, and the relationship between weather variables and the spread of the virus has been the subject of extensive research. Studies have shown that weather variables such as temperature, humidity, and rainfall can affect the transmission of COVID-19.

For instance, a study conducted in Bangladesh found that subtropical climate conditions, with a mean temperature of about 26.6°C, mean relative humidity of 64%, and rainfall of approximately 3 mm, were associated with increased COVID-19 cases from the onset. The same study found that for every 1 mm increase in rainfall, COVID-19 cases increased by 30.99%, while every 1°C increase of diurnal temperature was associated with a decrease of 14.2% in COVID-19 cases. In a review study conducted by McClymont and Hu, temperature was found to be a significant factor in the prevalence of COVID-19 cases in most studies, with the virus increasing as temperature decreases, and the highest incidence found within the range 0-17ºC. However, the relationship between temperature and COVID-19 cases is not consistent across all regions [36,37].

For instance, a study in Spain by Bashir found no consistent evidence of an association between temperature and COVID-19 cases. Similarly, a study in India by Goswami, Bharali, and Hazarika found a significant association between COVID-19 cases and temperature and relative humidity. However, this association was not consistent across all regions of India, and the differences in data available when the study was conducted may have influenced the variation in results [38,39].

Overall, the impact of weather on COVID-19 transmission and prevalence is complex and multifaceted, with variations in the relationship across different regions and populations. Nevertheless, understanding the influence of weather variables in COVID-19 transmission is vital for developing early warning systems, predicting seasonality, and informing public health policies aimed at controlling the spread of the virus.

Socio-Demographic Factors and COVID-19

The COVID-19 pandemic has affected individuals and communities differently across various socio-demographic factors. Studies have shown that age, income, education, race, employment, and gender are key factors that significantly impact COVID-19 outcomes. Age is a critical socio-demographic factor associated with COVID-19. Older adults are at higher risk of developing severe illness and death from COVID-19. According to the CDC, people over the age of 65 are at a higher risk of severe illness, hospitalization, and death from COVID-19 than younger age groups [26].

Rotejanaprasert found most of the cases of COVID-19 are among males between the ages of 20-49. However, COVID-19 related deaths are found mostly among the elderly aged 75 years and over [19]. Income and education levels also significantly impact COVID-19 outcomes. Individuals with lower income and education levels are more likely to have jobs that require them to work near others, making them more susceptible to COVID-19. Additionally, lower-income individuals may have less access to healthcare, testing, and other resources to prevent infection [40-42].

Race is another important socio-demographic factor associated with COVID-19 outcomes. Studies have shown that racial and ethnic minorities, such as Black, Hispanic, and Native American communities, have higher rates of COVID-19 infections, hospitalizations, and deaths than White communities . This may be due to systemic racism, healthcare disparities, and socioeconomic factors that disproportionately affect these communities. Employment is also a crucial factor associated with COVID-19 outcomes. Essential workers, such as healthcare workers, grocery store employees, and public transportation workers, are at higher risk of exposure to COVID-19 due to their job requirements [7-9].

Studies have shown that these workers are more likely to test positive for COVID-19 than individuals who work from home [42]. Gender is another socio-demographic factor associated with COVID-19 outcomes. Studies have shown that men are more likely to die from COVID-19 than women, although the reasons for this are not yet fully understood. This may be due to biological differences, such as hormone levels, or differences in behavior, such as adherence to preventive measures [43-45].

In conclusion, socio-demographic factors such as age, income, education, race, employment, and gender have a significant relationship with COVID-19 outcomes. Understanding these factors can help public health officials and policymakers develop targeted interventions to reduce the impact of the pandemic on vulnerable populations. Based on the literature review, understanding the various factors associated with COVID-19 cases and deaths is essential for developing effective interventions to control the spread of the virus. By considering the socioenvironmental, socioeconomic, weather, health, and sociodemographic factors, policymakers and health officials can create targeted interventions that address the specific needs of different communities.

For example, policies that aim to increase access to healthcare, testing, and other resources can help reduce the impact of COVID-19 on vulnerable populations, such as individuals with lower income and education levels. Policies

that target essential workers, such as providing personal protective equipment (PPE) or prioritizing vaccinations, can help reduce the risk of exposure and infection in these high-risk groups. Moreover, policies that address the broader social determinants of health, such as housing, employment, and education, can help reduce the underlying factors that contribute to health disparities and increase the risk of COVID-19 infection and death. In summary, a comprehensive approach that considers the various factors associated with COVID-19 cases and deaths is necessary for developing effective interventions to control the spread of the virus and mitigate its impact on vulnerable populations.

Methodology

Study Area

This research is focused on the state of Oklahoma in the US. As a result of the significant impact of COVID-19 in the state, with a significant number of virus-related cases and deaths, it is crucial to explore the impact of the pandemic on public health and healthcare systems in this region. As of February 22, 2023, when the data was still publicly released, the state has had a cumulative COVID-19 case count of 1,284,450, cumulative COVID-19 deaths of 17,887, weekly new cases of 2,229 and 318.4 average number of daily new cases . However, recent data from COVID-19 deaths as of January 29, 2024, reports that Oklahoma COVID-19 deaths tops 20,000, making it the state with the second highest COVID-19 deaths since 2020 [46,47].

Data Collection

To answer the research questions, secondary data from different sources were utilized. The COVID-19 data was collected from Oklahoma State Department of Health at the zip code level [46]. The weather data, which includes average temperature, humidity, and sun radiation, was collected from the Oklahoma Mesonet. The health data, which include disability, stroke, disability, diabetes, depression, BP-high, Kidney, arthritis, asthma, obesity, and cancer, was derived from CDC. The socio-economic and socio- demographic data such as poverty, unemployment, uneducated, Blacks, Hispanics and America Indians was collected from the US Census Bureau. The zip code boundary data shape file was downloaded from the US Census Bureau Tiger/line shapefile.

Data Analysis

The study employs two models for its analysis: The Multiscale Geographical Weighted Regression (MGWR) model and the Zero-inflated Negative Binomial (ZINB) regression model.

Multiscale Geographically Weighted Regression (MGWR)

Multiscale Geographically Weighted Regression (MGWR) is a spatial analysis technique that is used to model spatially varying relationships between a dependent variable and a set of independent variables. Unlike the traditional classic GWR model, MGWR model accounts for the difficulty of constraining the optimal bandwidth of all variables to be the same, by making it possible for variables to have varying spatial smoothing levels. The spatial process model produced by the MGWR multi-bandwidth approach is more realistic and practical since each variable's unique bandwidth may also be utilized as an index of the spatial size of each spatial process [48].

The MGWR is an improvement of the GWR model as it accounts for everything the GWR model stands for in addition to allowing each explanatory variable to operate on a different scale. This means that just like the GWR model, the MGWR model also accounts for spatial heterogeneity in the relationship between variables, allowing for the identification of local variations in the relationships . MGWR assumes that the relationship between the dependent variable and the independent variables varies across space [49-51].

In traditional regression techniques, it is assumed that the relationship between variables is constant across the entire study area. However, this assumption is often not valid, especially when dealing with spatial data where there may be significant spatial heterogeneity. MGWR addresses this issue by allowing the relationship between variables to vary across space. This is achieved by fitting a separate regression model for each location in the study area, using the local data within a specified radius or bandwidth. The result is a set of regression coefficients that are spatially varying, providing insights into the spatial variations in the relationships between variables [49,50].

One of the main advantages of MGWR is that it provides a more nuanced understanding of spatial relationships than traditional regression techniques. MGWR allows for the identification of spatially varying relationships, enabling more accurate predictions and better-informed policy decisions. Additionally, MGWR can be used to investigate the spatial distribution of the coefficients, providing insights into the spatial patterns of the relationship between variables. A major limitation of MGWR is that it assumes that the relationship between variables is linear, which may not always be the case.

Despite this limitation, it is a powerful spatial analysis technique that can provide insights into spatially varying relationships between variables. However, it should be used judiciously, considering the limitation of the method, and the suitability of the approach to the research question at hand. The MGWR and GWR models have been used in various studies related to COVID-19, providing useful insights into the spatial heterogeneity of the pandemic and its impact on different regions [51-53].

Therefore, in this study the MGWR model is also used to determine the geographical relationship between COVID-19 cases and potential socioeconomic, sociodemographic, weather and health factors such as age, temperature, race, employment, employment, diabetes, disability, obesity, education and income and the geographical heterogeneity presents in the relationship. The MGWR model is utilized to examine the geographical relationships between COVID-19 cases and potential risk factors. The formular for the MGWR is:

Equation 1: $y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + \epsilon_i$.

Where:

- y_i Dependent variable at location i.
- $\beta_0 (u_i, v_i)$ Intercept at location i with spatial coordinates (u_i, v_i).
- $\beta_k (u_i, v_i)$ Coefficient of the kth predictor at location i,which can vary from spatially.
- x_{ik} k-th independent variable at locations i.
- ϵ_i Random error term at location.
- P Number of independent variables.

Zero-Inflated Negative Binomial Regression (ZINB)

The zero inflated negative binomial model is a modification of the Poisson regression model, designed to deal with the problem of excess zero and overdispersion in count data. The ZINB model is employed to analyze the geographical relationships between COVID-19 deaths and potential risk factors. The model employs two separate processes for processing the data. The first is used to process the count values in the data, known as the count model coefficients, and the second is used to process the excess zeros, known as the zero-inflation model coefficient. Each is modeled independently.

The use of the zero inflated negative binomial model helps to address the issue of overdispersion in the data where the variance is much greater than the mean and where excess zeros exist in the dataset [54]. Using a non-zero inflated model in this case may lead to an underestimation of standard error and a potential overconfidence in the results. Hence, if count data consists of non-negative, highly skewed sequences count with a large proportion of zeros, the zero-inflated model is best for analyzing such data.

To further test for overdispersion in the data, an overdispersion test was carried out using R. In addition, to further test if the ZINB model is a good fit compared to a non-zero model (Negative Binomial (NB) a Vuong test was carried out using R. The formular for the ZINB is:

Equation 1: $\mathrm{P(Y = y|X, Z) = hfh} \begin{cases} \pi + (1-n).\left(\frac{\Gamma(r+y)}{\Gamma(r).y!}\right).\left(\frac{\mu^r}{(\mu+r)^{r+y}}\right), & if\ y = 0, \\ (1-\pi).\left(\frac{\Gamma\,(r+y)}{\Gamma(r).y!}\right).\left(\frac{\mu^r}{(\mu+r)^{r+y}}\right), & if\ y > 0, \end{cases}$

Where:

- Y Dependent variable (count data).
- π *probability of being in the* "excess zero" *group* (*modeled with a logit function*).
- μ Mean of the negative binomial distribution (linked to the predictors by a log link function).
- r dispersion parameter of the negative binomial distribution.
- Γ Gamma function.

Equation 2: $logit\ (\pi) = \ln\left(\frac{\pi}{1-\pi}\right) = Z^T\gamma$

Where:

- π *probability of being in the* "excess zero" s.
- Z Predictors for the zero-inflation process.
- γ *Coefficients for the zero-inflation preidctors.*

Equation 3: $L = \prod_{i=1}^{n}\left[\pi_i + (1 + \pi_i) f_{NB}(0|\mu_{i,}r\right]^{1_{yi=0}} . \left[(1-\pi_i) f_{NB}(y_i|\mu_{i,}r)\right]^{1_{yi>0}}$

Where:

- f_{NB} is the probability mass function of the negative binomial distribution.
- π *Explains structural* zeros i.e observations with inherent zero count.
- μ Explains the remaining *zeros and non- zeros counts based on a negative* binomial process.

Two distinct methods are used to analyze the COVID-19 cases and COVID-19 deaths due to significant differences in the data. The COVID-19 deaths data exhibits excess zeros, leading to overdispersion, while the COVID-19 cases data does not. Out of the 664 zip codes, the COVID-19 deaths had over 300 zip codes with zero value, while the COVID-19 cases

had only 4 zip codes with zeros. Therefore, the same methods cannot be applied for the analysis. Hence, a more robust model that could best address the excessive zeros in the data was used for the COVID-19 deaths.

In summary, the main objectives of this study are to examine the potential socioeconomic, sociodemographic, socioenvironmental, weather, and health factors that contribute to COVID-19 cases and deaths in Oklahoma and to analyze how these factors vary across different geographic locations. To achieve these objectives, the study utilizes MGWR model and the zero inflated negative binomial model to explore the local relationships between COVID-19 cases and deaths respectively and their driving factors.

Statistical Analysis Procedure

For this study, 20 independent variables were selected based on literature which includes; humidity, temperature, sun-radiation, 65 years and over, uneducated, America Indian, Black, Hispanics, employed, poverty, non-English speaking, stroke, depression, cancer, obesity, disability, asthma, arthritis, diabetes, BP-high, and kidney failure (Table 3.1). These independent variables were used to analyze the dependent variable COVID-19 case rate and death rate.

Variable	Description	Sources	Unit
Temperature	Average air temperature	Mesonet	Degree Fahrenheit
Humidity	Average relative humidity	Mesonet	Percentage
Sun radiation	Average sun radiation	Mesonet	Megajoules per meter square
Arthritis	Population with Arthritis	CDC	Percentage
BPHigh	Percentage of population with high BP	CDC	Percentage
Cancer	Population with cancer	CDC	Percentage
Casthma	Population with asthma	CDC	Percentage
Depression	population with depression	CDC	Percentage
Diabetes	Population with diabetes	CDC	Percentage
Kidney	Population with Kidney problem	CDC	Percentage
Obesity	Population with Obesity	CDC	Percentage
Stroke	population with stroke	CDC	Percentage
Disability	Population with disability	CDC	Percentage
Age 65 and over	Population age 65 and above	US Census Bureau	Percentage
Non-English speaking	Non-English-speaking population	US Census Bureau	Percentage
Poverty	Poverty population	US Census Bureau	Percentage
Uneducated	Less than high school degree	US Census Bureau	Percentage
Employed	Employed population	US Census Bureau	Percentage
Blacks	Black population	US Census Bureau	Percentage
American Indians	American Indians population	US Census Bureau	Percentage
Hispanics	Hispanic population	US Census Bureau	Percentage

Table 3.1: Independent Variables

Using SPSS software, the data were explored to check for normality in the variables as well as the residuals. To reduce the number of independent variables to a considerable size as well as account for the problem of multicollinearity, a factor analysis with a varimax rotation method was carried out (Table 3.1). Four factors were extracted which explains 68.8% of the variation in the independent variables. In addition to accounting for normality in the residual, the log using Log 10 was used for the COVID-19 cases data.

Groups	Variables	Factors 1	2	3	4
Weather	Humidity		0.86		
	Temperature				
	Solar radiation		-0.82		
Socio-economic characteristics	Uneducated			0.61	
	Employed			-0.69	
	Poverty			0.78	
Demographic characteristics	Age (65 years and above)				-0.60
	America Indians				
	Blacks			0.47	
	Hispanics				0.57
	Non-English Speaking				0.60
Health	Stroke	0.93			
	Depression	0.45			
	Cancer	0.83			
	Obesity	0.75			
	Disability	0.85			
	Asthma	0.63			
	Arthritis	0.93			
	Diabetes	0.90			
	BP-high	0.75			
	Kidney failure	0.96			
		Those with health conditions	Humidity and no solar radiation	Those with poor economic condition	Non-English-speaking young Hispanics

Table 3.2: Independent Variables with their Corresponding Highest Factor Loadings

Based on the factor loadings, four factors were selected and renamed according to their factor loadings. Factor loadings represent the correlation between the variable and the underlying factor. Factor loadings close to -1 or 1 indicate that the factor strongly influences the variable while factor loading close to 0 indicates a weak influence on the variable. The new factors are renamed based on their factor loading score as shown in Table 3.2.

When utilizing the MGWR methods, choosing the appropriate bandwidth is crucial because it determines the spatial scale at which each factor operates. MGWR allows different factors to operate at different spatial scales, so bandwidth selection needs to reflect this multiscale nature. There are two methods of selecting bandwidths for analysis, the fixed method, and the adaptive method. For the fixed method, a constant value is used for all the geographical areas, with the assumption that spatial influence is constant for all variables.

Whereas for adaptive method, bandwidth varies based on the data difference for each variable. The more the data the larger the bandwidths and vice versa. For this study, the adaptive bandwidth selection is used, and the ArcGIS software automatically selects different bandwidths for each factor by minimizing the Corrected Akaike Information Criterion (AICc), balancing model fit and complexity. The MGWR results produces a table that explains how each factor influences the result (Table 3.3)

FACTORS	SIG_ COUNT	SIG_PCT	NBR_ COUNT	NBR_ PCT	C_MEAN	C_STD	C_MIN	C_MAX	C_ MEDIAN
Intercept	1	0.16	53	8.26	0.03	0.27	-0.68	0.55	0.05
Those with health conditions	97	15.11	57	8.88	-0.01	0.32	-0.93	0.70	-0.00

Humidity and no solar radiation	226	35.20	122	19.00	0.32	0.19	-0.23	0.92	0.33
Those with poor economic condition	166	25.86	59	9.19	-0.06	0.39	-1.08	0.79	-0.03
Non-English-speaking young Hispanics	32	4.98	85	13.34	0.06	0.17	-0.49	0.57	0.07

Table 3.3: MGWR Results Summary

The table provides detailed explanations of how each factor behaves across different geographic locations, showing both the significance and the variability of the relationship in different regions. The SIG_COUNT column lists the number of locations where the variable has a statistically significant relationship with the dependent variable. For those with health conditions, 97 zip codes show a significant relationship. The SIG_PCT represents the percentage of total locations where the factors are statistically significant. For factor "Humidity and no solar radiation" the percentage is 35.20%, meaning that over 35% of the location showed significant relationships with this factor. The NBR_COUNT represents the number of neighbors that were used to estimate the local coefficients for the variable. The factor Humidity and no solar radiation has 122 neighbors used in the estimation process. The NBR_PCT represents the percentage of neighbors used to estimate the variable's local coefficients relative to the total zip codes available. For non-English-speaking young Hispanics, 13.34% of neighbors were used. The C_MEAN, C_STD, C_MIN, C_MAX, AND C_MEDIAN represent the mean, standard deviation, minimum, maximum, and median values of the local coefficients for the factors across all locations. The coefficient mean, standard deviation, minimum, maximum, and median for Humidity and no solar radiation are 0.32, 0.19, -0.23, 0.92, and 0.33 respectively. A 0.32 mean value indicates a positive relationship with the dependent variable. However, its influence varies across locations, as shown by the range between its minimum value (-0.23) and maximum value (0.33).

Results

COVID-19 Cases Analysis Results

For the COVID-19 cases, the four transformed independent factors were utilized, which include those with health conditions, humidity and no solar radiation, Those with poor economic conditions and non-English-speaking young Hispanics. The results as shown in Figure 3.1 suggest that only areas with color green and white can be used to determine the significance of the variables given their R square value of at least 0.5. In general, the model seems to do very well with an overall R2 of 0.58. In addition, there seems to be no problem of non-normality and heteroscedasticity (Figure 3.2). However, although all variables were significant, not all variables were significant in all locations (Figures 3.3 through 3.6).

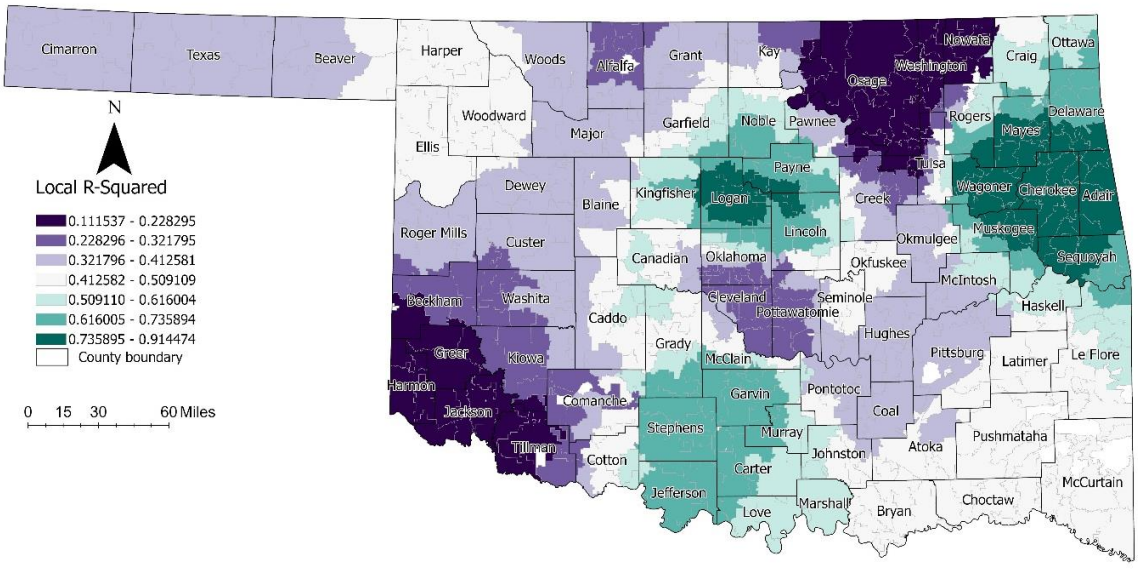

Figure 3.1: The R Squared for COVID-19 Cases

(a) (b)

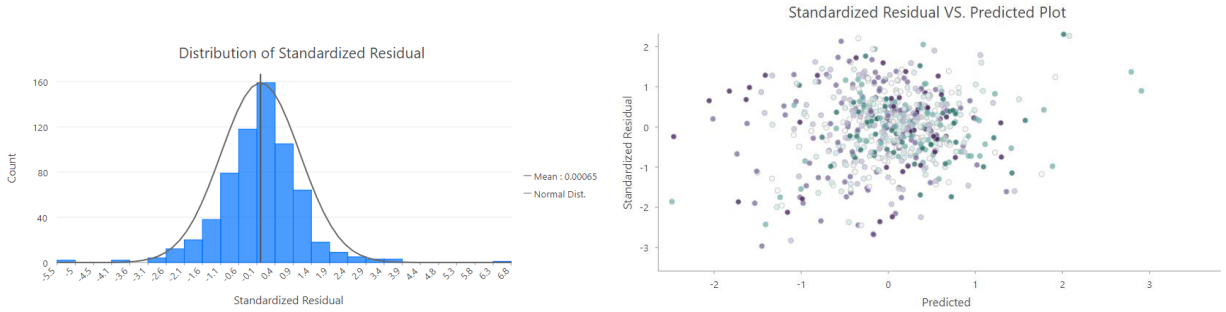

Figure 3.2: The Residuals Plot for COVID-19 Cases: (a) Histogram (b) Scatter plots

The first variable, "those with health conditions," was only significant in some zip codes with R square above 0.5. Zip codes in Logan, Payne, Garfield, Lincoln and Washita counties showed a significant positive correlation whereas zip codes around Delaware, Mayes, Wagoner, Cherokee and Adair showed a negative significant relationship with COVID-19 based on the significance and coefficient map (Figure 3.3). This indicates that a one standard deviation increase in those with health conditions is correlated with a corresponding increase between 0.3 to 1.0 units in COVID-19 cases in zip codes such as Payne and between 0.3 to 1.0 units decrease in zip code around counties such as Delaware and Mayes (Figure 3.3).

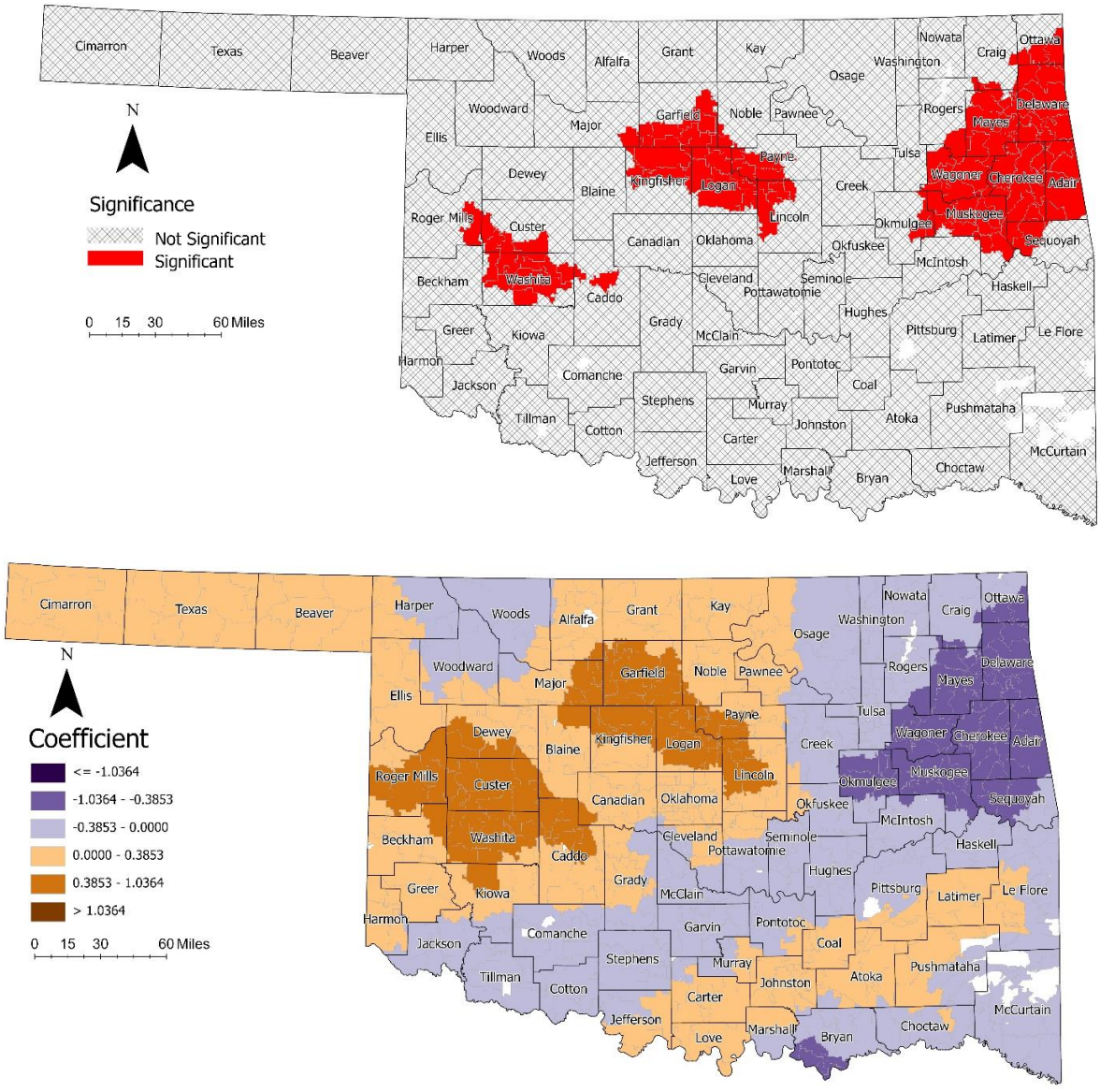

Figure 3.3: The (a) Significance and (b) Coefficient Maps of the Variable "Those with Health Conditions" in the Model

The second variable, "humidity and no solar radiation," was significant in some zip codes with at least 0.5 R square. Zip codes in Deware Cheroke, Adair, Sequoya, Leflore, Pushmataha, Choctaw, McCurtain, Stephens, Jefferson, Payne, Noble, Harper, Woodwod, and Ellis Counties all had positive correlation with COVID-19 cases (Figure 3.4). This means that one standard deviation increases in humidity and no solar radiation is significantly correlated with increase in COVID-19 cases between 0.3 to 1.0 units in the corresponding zip codes.

The third variable, "those with poor economic condition" was only significant in some zip codes with at least 0.5 R square, with both positive and negative correlation. Zip codes in Ottawa, Craig, Delaware, Mayes, Cherokee and Adair, Sequoyah, Harper, Elis and Woodward Counties and all had positive correlation with COVID-19 cases. Whereas zip codes in Leflore, McCurtain, Pushmataha, Latimer, Stephens, Jefferson, Love, Garvin, and Carter had a negative significant correlation with COVID-19 cases (Figure 3.5).

This means that one standard deviation increases in humidity and no solar radiation is significantly correlated with increase in COVID-19 cases between 0.3 to 1.0 units in Ottawa, Craig, Delaware, Mayes, Cherokee and Adair, Sequoyah, Harper, Elis and Woodward Counties and decrease between 0.3 to 1.0 in Leflore, McCurtain, Pushmataha, Latimer, Stephens, Jefferson, Love, Garvin, and Carter.

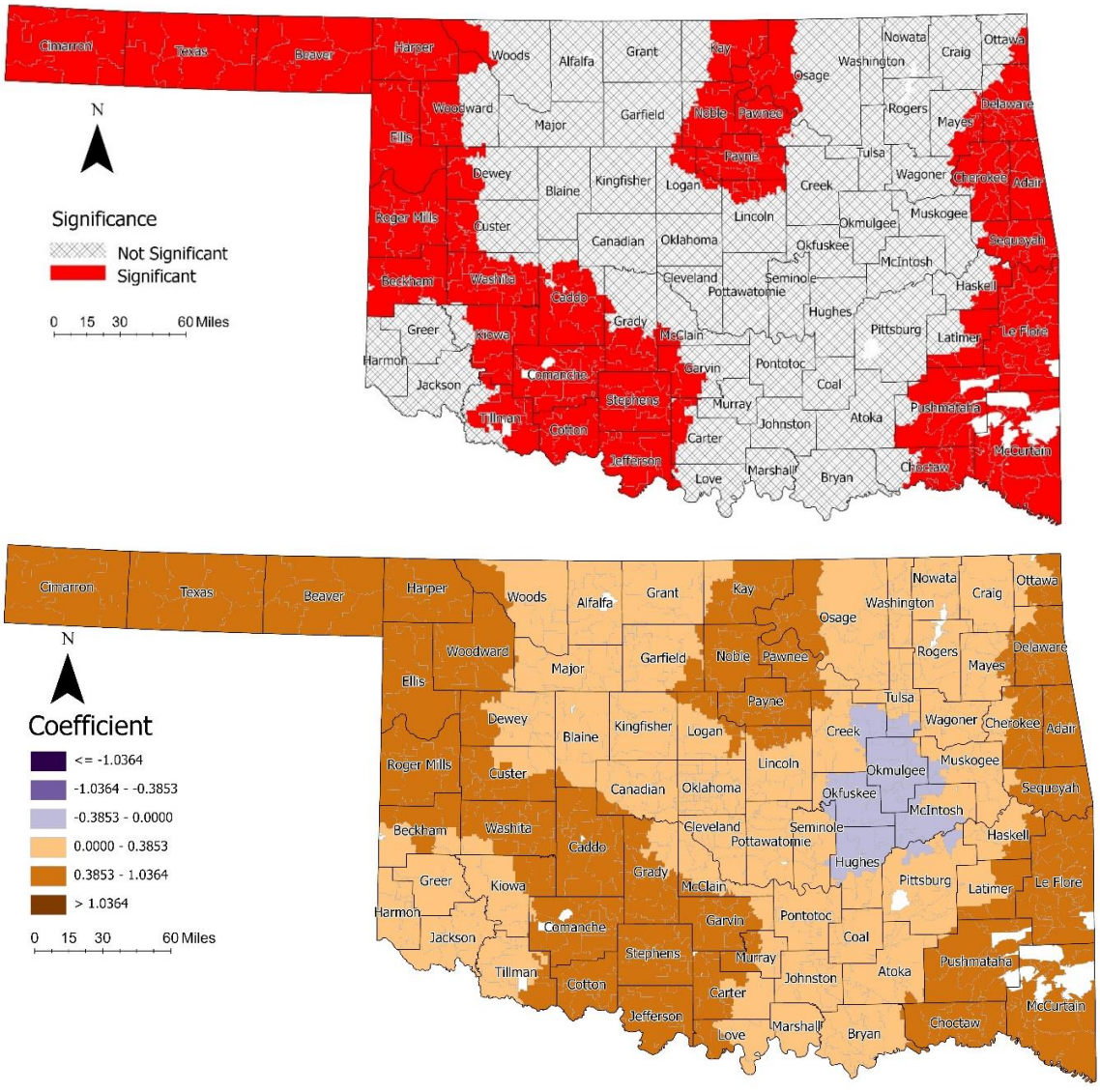

Figure 3.4: The (a) Significance and (b) Coefficient Maps of the Variable "Humidity and no Solar Radiation" in the Model

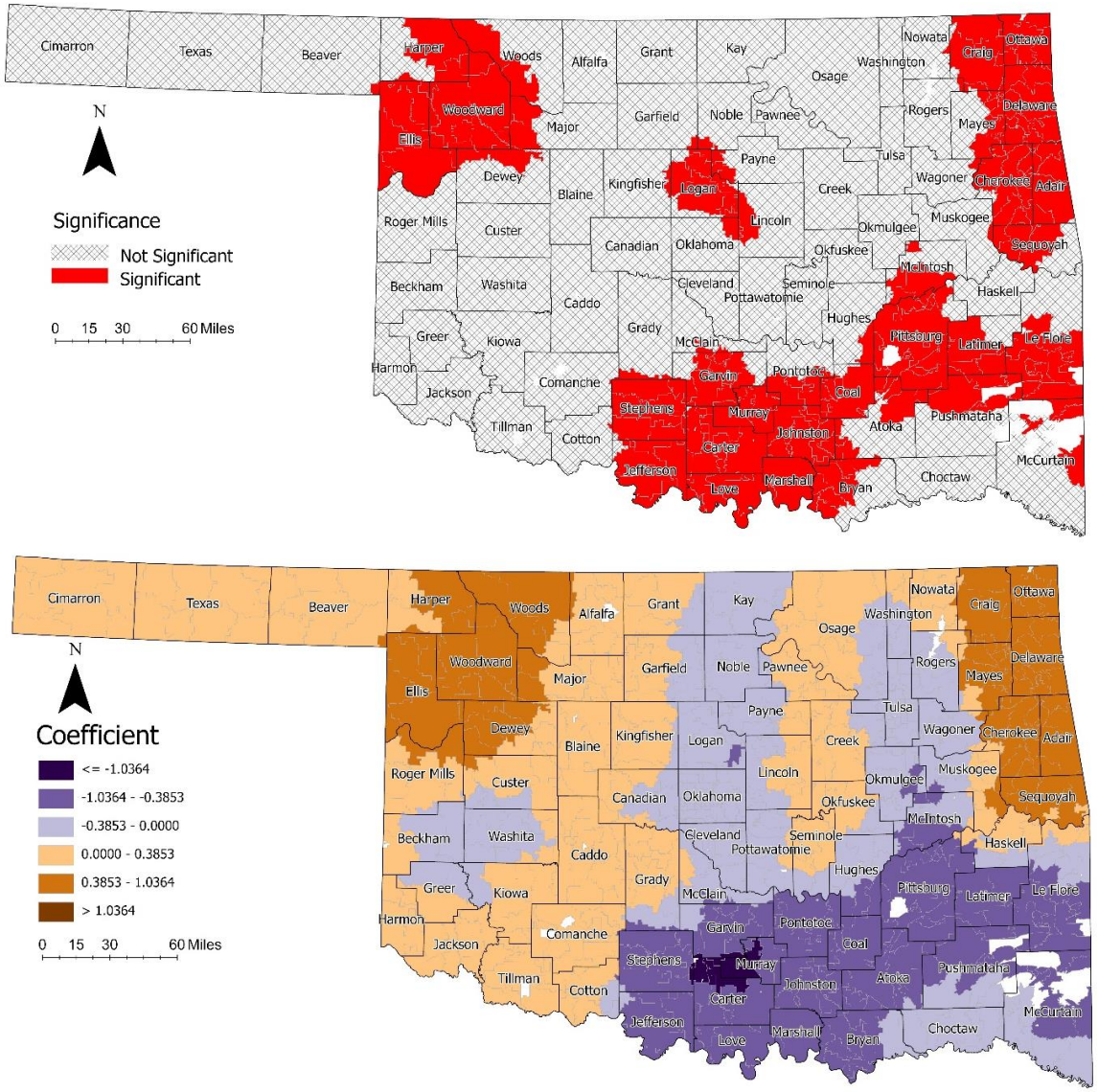

Figure 3.5: The (a) Significance and (b) Coefficient Maps of the Variable those with Poor Economic Condition" in the model

The fourth variable, "non-English speaking, young Hispanics," was only significant in some zip with at least 0.5 R square. Zip codes in Bryan had positive correlation whereas zip code in Cherokee, Adair, and Sequoyah had negative correlation (Figure 3.6). This indicates that a one standard deviation increase non-English speaking young Hispanics is significantly correlated with increase in COVID-19 cases between 0.3 to 1.0 units in Bryan County and decrease between 0.3 to 1.0 units in Cherokee, Adair, and Sequoyah zip codes.

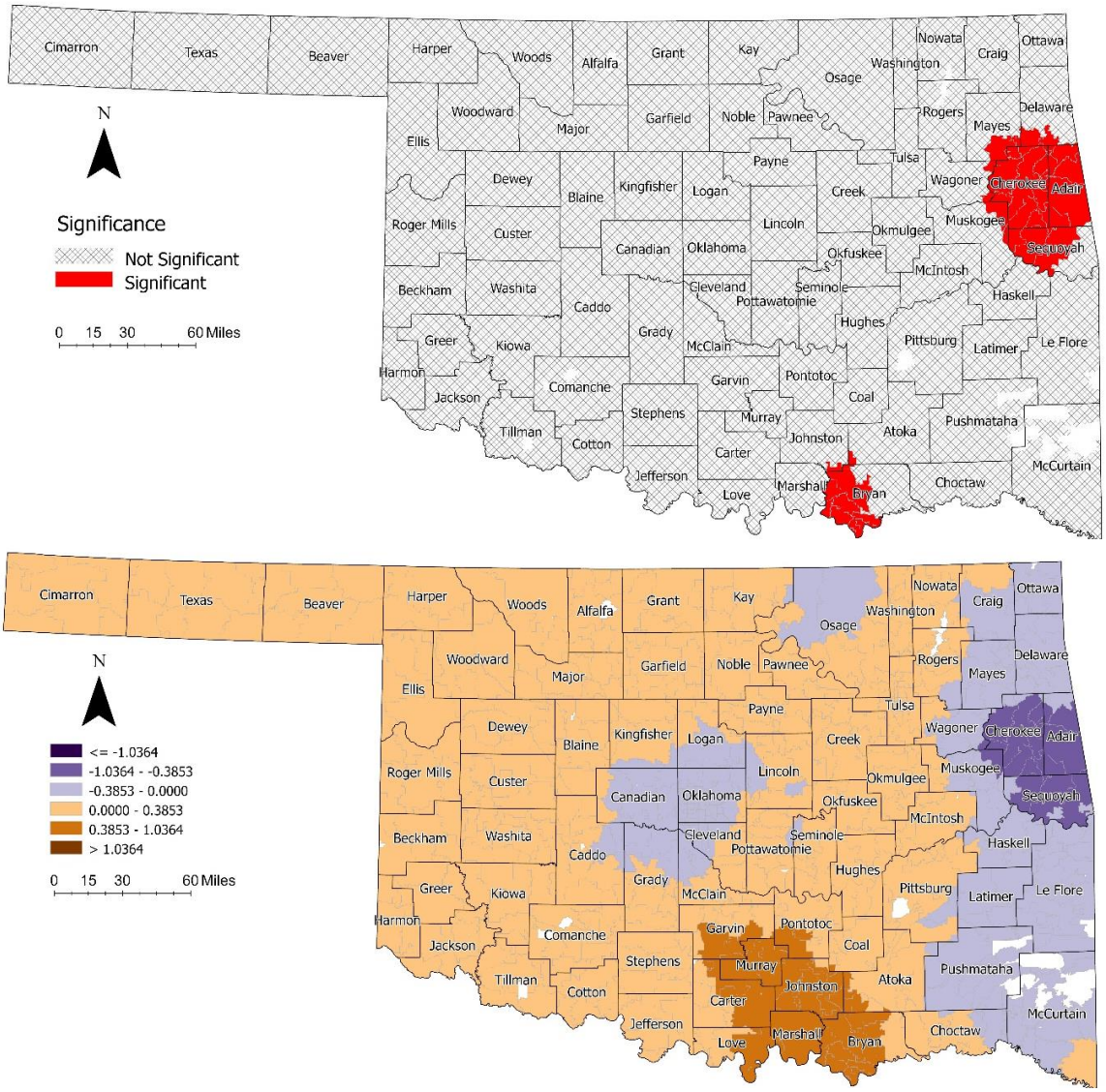

Figure 3.6: The (a) Significance and (b) Coefficient Maps of the Variable "Non-English Speaking Young Hispanics" in the Model

The results described above highlight the importance and capability of MGWR in accurately estimating the coefficients of a local regression model, demonstrating that the significance of explanatory variables varies across different regions within the same study area. The observation of both positive and negative relationships between certain factors, such as "those with health conditions, those with poor economic condition, and non-English speaking young Hispanics and COVID-19 cases in some areas can be attributed to differences in socioeconomic, demographic, and policy contexts.

In regions where the population has better access to healthcare services, higher income, lower population density, public interventions, community outreach, and education, the relationship might be negative, as these factors could have mitigated the impact of the pandemic. For instance, Adair County in Oklahoma, although associated with low income, is also known for its low population density, which may have contributed to the negative association between those with health conditions and COVID-19 cases in Adair County. Similarly, Logan County, which showed a negative relationship between those with poor economic condition and COVID-19, has a relatively rural population and known for its high income, which might explain the negative association observed in the county. Therefore, the observation of both positive and negative relationships between certain factors, such as "those with health conditions, those with poor economic condition, non-English speaking young Hispanics and COVID-19 cases can be attributed to differences in the sociodemographic and economic distributions across these areas, thus explaining the varied associations observed.

COVID-19 Deaths Analysis Results

For the COVID-19 deaths, the same four factors used for the COVID-19 cases were employed for the analysis. The ZINB was used to model the COVID-19 deaths using the COVID-19 death rate as the dependent variable and the four factors as the independent variables. As shown in Figure 3.7, over 300 zip codes recorded no COVID-19 related deaths. In addition, the overdispersion test shown in Table 3.4 below has a P-value less than 0.05%, indicating that the data is overly dispersed. Both results therefore make the ZINB model regression a better model for the analysis.

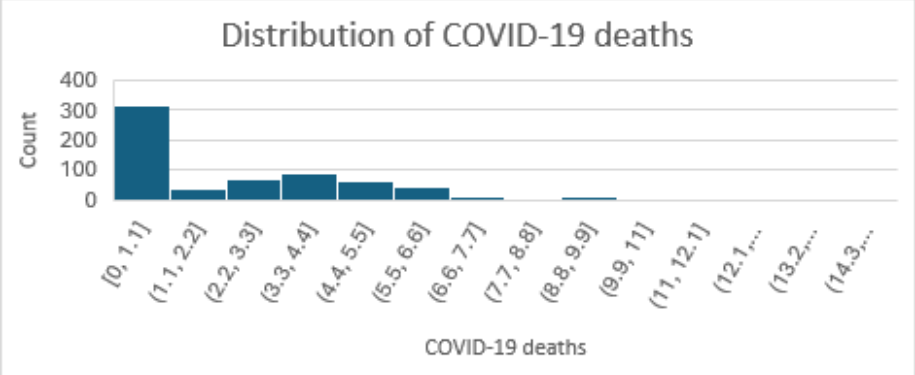

Figure 3.10: COVID-19 Deaths Distribution Data

Statistic	Value
z	7.5531
p-value	2.125e-14
Alternative Hypothesis	True dispersion is greater than 1
Sample Estimate	Dispersion = 3.254212.

Table 3.4: Overdispersion Test

The ZINB model processes two separate results, the count model and the zero inflated model. The result of the count model as shown in Table 3.5 illustrates that only 3 variables play a significant role in COVID-19 deaths. These variables are, "those with health conditions," "humidity and no solar radiation," and "those with poor economic condition." Those with health conditions with a coefficient of 0.4 and a positive significant relationship suggesting that the more people with health conditions the more likely COVID-19 deaths will increase. In contrast, for the variable 'humidity and no solar radiation," and "those with poor economic condition" there is a negative relationship which denotes that the less humidity and no solar radiation and the less of those with poor economic condition the more the likelihood of COVID-19 deaths.

For the zero-inflation model used to predict excess zeros, the results suggest that all variables were significant (Table 3.5). The variable "those with health conditions," results was positively significant which suggests that the more people with health conditions the less likely there will be an excess zero in the model. The other variables were negatively significant which are "humidity and no solar radiation," "Those with poor economic condition", and "non-English speaking young Hispanics" suggest that the less of these variables the less likely there will be an excess zero in the model. In addition, to test the validity of the ZINB model, a Vuong test was conducted given that the model is non-nested. The results as shown in Table 3.6 with a P-value less than 0.05% indicate that the ZINB model is the most suitable for illustrating the association between COVID-19 deaths and the independent variables used for the study.

	Negative Binomial		Zero-inflation (binomial with log link)		Count model coefficients (poisson with log link)	
	Marginal mean		Marginal mean		Marginal mean	
Variable	Est	Se	Est	Se	Est	Se
Intercept	1.51	<2e-16***	-0.22	0.02*	0.81	<2e-16***
Those with health conditions	0.41	8.74e-14***	0.55	0.001**	0.18	1.52e-06***
Humidity and no solar radiation	-0.10	0.00214**	-0.39	0.00017***	0.10	0.00016***
Those with poor economic condition	-0.13	0.00136**	-0.29	0.03*	0.02	0.53
Non-English-speaking young Hispanics	0.06	0.011	-0.88	3e-10***	0.22	5.89e-13***

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Table 3.5: Zero Inflated Negative Binomial

Statistic	Value	Alternative Hypothesis	p-value
aw	-16.63750	ZINB > NB	< 2.22e-16
AIC-corrected	-16.50300	ZINB > NB	< 2.22e-16
BIC-corrected	-16.20296	ZINB > NB	< 2.22e-16

Table 3.6: Vuong Z Statistic

Discussion

The study focuses on identifying key factors influencing the likelihood of COVID-19 cases and deaths in Oklahoma, using count data collected from the Oklahoma State Department of Health from January 2020 to October 2023. I used two separate models to analyze the data, MGWR and ZINB. I employed The MGWR to analyze the COVID-19 cases data, and the ZINB model was used for the COVID-19 deaths as it best addressed the excess zeros and overdispersion in the data.

The result of the MGWR model used for the COVID-19 cases suggests that three independent variables used for the analysis had both negative and positive significant relationships with the COVID-19-cases in various locations. The results as shown in Figures 3.3-3.6 tend to conform with literature, given that all four variables have strong explanatory powers to the increase and decrease of COVID-19 cases. The variable "those with health conditions" showed both positive and negative correlation with COVID-19 cases at various locations (Figure 3.3).

This means that an increase in people with health conditions may not necessarily lead to an increase in COVID-19 cases in all locations. Kompaniyets , in the early stage of the virus, found that 94.9% of patients hospitalized with COVID-19 had at least one underlying health condition [54]. Other studies also suggest that those with underlying conditions tend to have a significant positive association with COVID-19. However, based on this study, the result suggests that some locations tend to have a negative significant correlation with COVID-19. This means that having underlying conditions may not necessarily mean one is susceptible to COVID-19. In this case other location factors such as income and other social factors may come to play. This finding challenges the assumption that those with underlying conditions are more susceptible to COVID-19 virus, other factors may mitigate this risk [55-57].

For instance, higher income locations may practice healthier lifestyles and have better access to healthcare, reducing the overall impact of COVID-19 in those areas compared to low-income areas that might be faced with greater economic challenges, and limited healthcare access limiting increasing their susceptibility. For example, zip codes in Adair County with a negative significant association between underlying health conditions and COVID-19 cases are characterized by low income but also low population density. Low income might typically be associated with limited access to healthcare, which could exacerbate health conditions and increase vulnerability to severe outcomes from COVID-19.

However, the low population density in these areas plays a critical role in mitigating the spread of the virus. With fewer interactions between individuals and reduced opportunities for close contact, the transmission rate of COVID-19 may be significantly lower. This reduced transmission can result in fewer reported cases, despite the presence of underlying health conditions. Thus, the unique sociodemographic and geographic context of Adair County, specifically the combination of low income and low population density, helps explain the observed negative association. The second variable, "humidity and no solar radiation," showed a positive significant relationship with COVID-19 (Figure 3.4).

This means that an increase in humidity and absence of solar radiation would lead to an increase in COVID-19 cases. The result is in conformity with some past research. Although some studies have found an increase in humidity leading to decrease in COVID-19 cases, some have equally found that increase in humidity will not necessarily lead to a decline in COVID-19 cases. Hence, the role of location is important in better understanding and mitigating infectious diseases such as COVID-19 as what works for location A (for example) may not work for location B due to other environmental factors that may come to play [58-60].

The third variable "those with poor economic condition" showed both positive and negative significant correlation with COVID-19 cases (Figure 3.7). This means that some locations are susceptible to COVID-19 cases with the increase in those with poor economic condition and some locations don't. The result is in conformity with literature as studies have shown that the impact of income and employment as regards COVID-19 varies across states. For example, Logan County, which exhibited a negative relationship between those with poor economic condition and COVID-19 cases, demonstrates how unique socioeconomic factors can influence pandemic outcomes [61].

Despite having an average population density, the county is recognized for its high-income levels, which play a crucial role in mitigating the impact of poverty, low education and unemployment on COVID-19 transmission and case rates. High-income households may have better access to healthcare, testing, vaccinations, and other preventative measures, which can reduce the transmission and severity of COVID-19 cases in the region.

The fourth variable, "non-English speaking young Hispanics" showed both positive and negative significant correlation with COVID-19 cases (Figure 3.5). This means that increase in non-English speaking young Hispanics may not lead to an increase in COVID-19 cases in all locations as found in this study. This again may be due to several other social factors and environmental differences. Studies have shown that there is a greater chance of non-English speakers to test positive for COVID-19 even after adjusting for social factors, age, and ethnicity. However, the negative significant association between non-English variables and COVID-19 cases in Adair County which is characterized by low income and low population density, suggests that the unique sociodemographic and geographic factors in the area might be mitigating the spread of the virus among this group [62].

Low population density naturally reduces the frequency of interpersonal interactions, limiting opportunities for the virus

to spread. In rural areas like Adair County, the physical separation of households and communities plays a critical role in reducing COVID-19 transmission, including among non-English-speaking young Hispanics populations. The result of the COVID-19 cases and the explanatory variables using the MGWR model emphasize the heterogeneity of the associations in various locations, thereby emphasizing the importance of local/community mitigation policies as a better strategy to address infectious disease such as COVID-19. For the COVID-19 deaths, the same explanatory factors (those with health conditions, humidity and no solar radiation, those with poor economic condition, and non-English speaking young Hispanics) using the ZINB model statistical technique were utilized for the analysis.

The result suggests that only three variables, "those with health conditions," "humidity and no solar radiation," and "those with poor economic condition" had significant correlations with COVID-19 deaths. The variable, "those with health conditions" showed a positive significant relationship with COVID-19 deaths (Table 3.2).

This suggests that the increase in those with health condition will lead to an increase in COVID-19 deaths. The result is in conformity with most research which suggests that the mortality rate of COVID-19 patients with underlying health conditions is several times higher than that of patients without underlying health conditions. The result therefore suggests that more attention should be paid to those with underlying health conditions when creating policies to address and mitigate the proliferation of infectious diseases such as COVID-19. In addition, there is need for unprecedented commitment by the government and health authorities to all aspects of prevention, vaccination, and public health [63].

The variable "humidity and no solar radiation" had a negative significant correlation with COVID-19 deaths (Table 3.3). This means that the increase in humidity and no solar radiation results in a decrease in COVID-19 deaths [59]. The result is in conformity with most research as recent studies have found that the COVID-19 pandemic may be partially controlled with increase in temperature [64,65].

The last significant variable, "those with poor economic condition," showed a negative significant correlation with COVID-19 deaths. This means the increase in those with poor economic condition may decrease COVID-19 deaths. Recent studies tend to suggest a great national decrease in COVID-19 deaths . In Oklahoma, the increase in COVID-19 vaccination rates is a key factor contributing to the decrease in COVID-19 deaths among the populations in the state. Vaccination has been one of the most effective tools for reducing severe outcomes and mortality associated with COVID-19, and efforts to ensure equitable vaccine distribution may have played a pivotal role in mitigating the disproportionate impact of the pandemic on poor communities [66].

As of May 2023, COVID-19 vaccination rates in Oklahoma have shown considerable progress, though disparities among racial and ethnic groups persist. According to the Oklahoma State Department of Health's Weekly Epidemiology Report, approximately 74.3% of the state's population had received at least one dose of a COVID-19 vaccine, with 60.2% fully vaccinated. The increase in COVID-19 vaccination rates among populations in Oklahoma may have significantly contributed to the observed negative association between COVID-19 deaths and those with poor economic conditions in the state. Vaccination has been a critical factor in reducing severe outcomes and mortality, and targeted efforts to improve vaccine uptake in historically underserved communities may play a vital role in this outcome.

Based on the results of this study, it is evident that several underlying factors influencing COVID-19, as discussed above, are consistent with the literature. However, certain aspects, such as the heterogeneity in COVID-19 cases require further study and may provide basis for future research. Policymakers and government officials should consider locational differences when creating policies to address COVID-19 and infectious diseases in general. Tailoring policies to local contexts rather than applying global policies uniformly can lead to more effective outcomes.

Strengths and Limitations

The study utilizes quality up-to-date zip code level data from January 2020 through October 2023 collected from the Oklahoma State Department of Health. The use of up-to-date zip code level data is a strength as it provides the most current information about COVID-19 in the state. However, it also offers a major limitation, which is its applicability is limited to Oklahoma only. Regardless, similar methods can be used in other places of the world to determine key factors influencing infectious diseases such as COVID-19. Another limitation of the study is the aggregation of COVID-19 cases and deaths over a three-year period. Analyzing the data separately for different phases of the pandemic could yield varying results and lead to different conclusions. In addition, as the pandemic nears its end, reported cases are becoming less reliable. Consequently, data accuracy and reliability present another significant challenge. Also, there were unidentifiable zip codes due to wrong address as noted by OSDH. However, data collected from OSDH included both estimated and actual cases [67].

Conclusions

The study examines several underlying factors influencing COVID-19 in Oklahoma. To answer the first research question, "what social risk factors are associated with COVID-19 cases in Oklahoma", the MGWR method was used. The result indicates that all four explanatory variables (those with health conditions, humidity and no solar radiation, those with poor economic condition, and non-English speaking young Hispanics) used for the study had significant correlation with COVID-19 cases in different locations (Figures 3.3-3.6) [68-70].

However, for some of the variables there was heterogeneity in their associations at various locations, which has been determined to be caused by socio economic and demographic differences in the regions. To answer the second research question, "what social risk factors are associated with COVID-19 deaths in Oklahoma", the ZINB model was used due to the excess zeros and over dispersion in the data. The same explanatory variables as the COVID-19 cases were used for the COVID-19 deaths. The results suggest that only three variables (those with health conditions, humidity and no solar radiation, and those with poor economic condition) had significant association with COVID-19 deaths (Table 3.4) [71-72].

This study provides a useful guide to health care providers and policy makers in understanding regional differences when making policies and providing healthcare services and the need to apply local policies as a better strategy to address issues of infectious diseases such as COVID-19 in the state of Oklahoma.

References

1. Van Dorn, A., Cooney, R. E., & Sabin, M. L. (2020). COVID-19 exacerbating inequalities in the US. *Lancet (London, England), 395*(10232), 1243.
2. CDC 2023a. (n.d.). CDC COVID Data tracker. Centers for Disease Control and Prevention. Modified February 23, 2023. Retrieved March 2, 2023.
3. CDC 2024. (n.d.). Covid Data tracker. Centers for Disease Control and Prevention. Retrieved April 15, 2024.
4. O'Neill, M. S., Zanobetti, A., & Schwartz, J. (2005). Disparities by race in heat-related mortality in four US cities: the role of air conditioning prevalence. *Journal of urban health, 82,* 191-197.
5. CDC 2023b. (n.d.). People with certain medical conditions. Centers for Disease Control and Prevention. Modified February 26, 2023. *Retrieved February 27,* 2023.
6. Yanez, N. D., Weiss, N. S., Romand, J. A., & Treggiari, M. M. (2020). COVID-19 mortality risk for older men and women. *BMC public health, 20*(1), 1742.
7. Khanijahani, A., Iezadi, S., Gholipour, K., Azami-Aghdash, S. and Naghibi, D., 2021. A systematic review of racial/ethnic and socioeconomic disparities in COVID-19. *International Journal for Equity in Health, 20*(1):1-30.
8. Magesh, S., John, D., Li, W. T., Li, Y., Mattingly-App, A., Jain, S., ... & Ongkeko, W. M. (2021). Disparities in COVID-19 outcomes by race, ethnicity, and socioeconomic status: a systematic review and meta-analysis. *JAMA network open, 4*(11), e2134147-e2134147.
9. Mude, W., Oguoma, V. M., Nyanhanda, T., Mwanri, L., & Njue, C. (2021). Racial disparities in COVID-19 pandemic cases, hospitalisations, and deaths: A systematic review and meta-analysis. *Journal of global health, 11,* 05015.
10. Reyes, M. V. (2020). The disproportional impact of COVID-19 on African Americans. *Health and human rights, 22*(2), 299.
11. Sze, S., Pan, D., Nevill, C. R., Gray, L. J., Martin, C. A., Nazareth, J., ... & Pareek, M. (2020). Ethnicity and clinical outcomes in COVID-19: a systematic review and meta-analysis. *EClinicalMedicine, 29.*
12. Shakespeare, T., Ndagire, F., & Seketi, Q. E. (2021). Triple jeopardy: disabled people and the COVID-19 pandemic. *The Lancet, 397*(10282), 1331-1333.
13. Bwire, G. M. (2020). Coronavirus: why men are more vulnerable to Covid-19 than women?. *SN comprehensive clinical medicine, 2*(7), 874-876.
14. Fabião, J., Sassi, B., Pedrollo, E. F., Gerchman, F., Kramer, C. K., Leitão, C. B., & Pinto, L. C. (2022). Why do men have worse COVID-19-related outcomes? A systematic review and meta-analysis with sex adjusted for age. *Brazilian Journal of Medical and Biological Research, 55*, e11711.
15. De La Vega, R., Ruíz-Barquín, R., Boros, S., & Szabo, A. (2020). Could attitudes toward COVID-19 in Spain render men more vulnerable than women?. *Global public health, 15*(9), 1278-1291.
16. Gilmore, B., Ndejjo, R., Tchetchia, A., De Claro, V., Mago, E., Lopes, C., & Bhattacharyya, S. (2020). Community engagement for COVID-19 prevention and control: a rapid evidence synthesis. *BMJ global health, 5*(10), e003188.
17. Volkin, S., Mayer, R. E., & Dingwall, A. (2020). Let COVID-19 Serve as a Catalyst to Fix National Crisis of Poor Maternal Mortality Data. *Journal of Public Health Management and Practice, 26*(6), 525-527.
18. Luo, Y., Yan, J., & McClure, S. (2021). Distribution of the environmental and socioeconomic risk factors on COVID-19 death rate across continental USA: a spatial nonlinear analysis. *Environmental Science and Pollution Research, 28*(6), 6587-6599.
19. Rotejanaprasert, C., Lawpoolsri, S., Pan-Ngum, W., & Maude, R. J. (2020). Preliminary estimation of temporal and spatiotemporal dynamic measures of COVID-19 transmission in Thailand. *PloS one, 15*(9), e0239645.
20. Mena, G. E., Martinez, P. P., Mahmud, A. S., Marquet, P. A., Buckee, C. O., & Santillana, M. (2021). Socioeconomic status determines COVID-19 incidence and related mortality in Santiago, Chile. *Science, 372*(6545), eabg5298.
21. Sharma, A. K., Gupta, R., Baig, V. N., Singh, T. V., Chakraborty, S., Sunda, J. P., ... & Katoch, V. M. (2021). Socioeconomic status and COVID-19 related outcomes in India: hospital based study. *medRxiv, 2021-05.*
22. Sharma, A. K., Gupta, R., Baig, V. N., Singh, V. T., Chakraborty, S., Sunda, J. P., ... & Katoch, V. M. (2022). Educational status and COVID-19 related outcomes in India: hospital-based cross-sectional study. *BMJ open, 12*(2), e055403.
23. Yancy, C. W. (2020). COVID-19 and african americans. *Jama, 323*(19), 1891-1892.
24. Garg, S. (2020). Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed

coronavirus disease 2019—COVID-NET, 14 States, March 1–30, 2020. MMWR. *Morbidity and mortality weekly report, 69.*
25. Guasti, N. (2020). The plight of essential workers during the COVID-19 pandemic. *Lancet, 395*(10237), 1587.
26. Weaver, A. K., Head, J. R., Gould, C. F., Carlton, E. J., & Remais, J. V. (2022). Environmental factors influencing COVID-19 incidence and severity. *Annual review of public health, 43*(1), 271-291.
27. CDC 2022. (n.d.). What is health equity? Centers for Disease Control and Prevention. Modified July 1 2022. March 1, 2023.
28. Ejaz, H., Alsrhani, A., Zafar, A., Javed, H., Junaid, K., Abdalla, A. E., ... & Younas, S. (2020). COVID-19 and comorbidities: Deleterious impact on infected patients. *Journal of infection and public health, 13*(12), 1833-1839.
29. Guo, W., Li, M., Dong, Y., Zhou, H., Zhang, Z., Tian, C., ... & Hu, D. (2020). Diabetes is a risk factor for the progression and prognosis of COVID-19. *Diabetes/metabolism research and reviews, 36*(7), e3319.
30. Mayo Clinic 2022. COVID-19: Who's at higher risk of serious symptoms? Mayo Clinic. *Modified September 27 2022. Retrieved February 27,* 2023.
31. Sanyaolu, A., Okorie, C., Marinkovic, A., Patidar, R., Younis, K., Desai, P., ... & Altaf, M. (2020). Comorbidity and its impact on patients with COVID-19. *SN comprehensive clinical medicine, 2*, 1069-1076.
32. Petrakis, D., Margină, D., Tsarouhas, K., Tekos, F., Stan, M., Nikitovic, D., ... & Tsatsakis, A. (2020). Obesity-a risk factor for increased COVID-19 prevalence, severity and lethality. *Molecular medicine reports, 22*(1), 9-19.
33. Gregory, J. M., Slaughter, J. C., Duffus, S. H., Smith, T. J., LeStourgeon, L. M., Jaser, S. S., ... & Moore, D. J. (2021). COVID-19 severity is tripled in the diabetes community: a prospective analysis of the pandemic's impact in type 1 and type 2 diabetes. *Diabetes care, 44*(2), 526-532.
34. Richardson, S., Hirsch, J. S., Narasimhan, M., Crawford, J. M., McGinn, T., Davidson, K. W., ... & Northwell COVID-19 Research Consortium. (2020). Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *jama, 323*(20), 2052-2059.
35. Nagaveni, S., Vashishth, A., & Singh, B. Applications, Challenges and Health issues in Diabetes Management for COVID-19 Pandemic.
36. Islam, A. R. M. T., Hasanuzzaman, M., Shammi, M., Salam, R., Bodrud-Doza, M., Rahman, M. M., ... & Huq, S. (2021). Are meteorological factors enhancing COVID-19 transmission in Bangladesh? Novel findings from a compound Poisson generalized linear modeling approach. *Environmental Science and Pollution Research, 28,* 11245-11258.
37. McClymont, H., & Hu, W. (2021). Weather variability and COVID-19 transmission: A review of recent research. *International journal of environmental research and public health, 18*(2), 396.
38. Bashir, M. F., Jiang, B., Komal, B., Bashir, M. A., Farooq, T. H., Iqbal, N., & Bashir, M. (2020). Correlation between environmental pollution indicators and COVID-19 pandemic: a brief study in Californian context. *Environmental research, 187*, 109652.
39. Goswami, K., Bharali, S., & Hazarika, J. (2020). Projections for COVID-19 pandemic in India and effect of temperature and humidity. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 14*(5), 801-805.
40. Sartorius, B., Lawson, A. B., & Pullan, R. L. (2021). Modelling and predicting the spatio-temporal spread of COVID-19, associated deaths and impact of key risk factors in England. *Scientific reports, 11*(1), 5378.
41. Hall, L. R., Sanchez, K., da Graca, B., Bennett, M. M., Powers, M., & Warren, A. M. (2022). Income differences and COVID-19: Impact on daily life and mental health. *Population Health Management, 25*(3), 384-391.
42. Kim, S. Y., Yoo, D. M., Min, C., & Choi, H. G. (2021). The effects of income level on susceptibility to COVID-19 and COVID-19 morbidity/mortality: a nationwide cohort study in South Korea. *Journal of Clinical Medicine, 10*(20), 4733.
43. Gaitens, J., Condon, M., Fernandes, E., & McDiarmid, M. (2021). COVID-19 and essential workers: a narrative review of health outcomes and moral injury. *International Journal of Environmental Research and Public Health, 18*(4), 1446.
44. Kharroubi, S. A., & Diab-El-Harake, M. (2022). Sex-differences in COVID-19 diagnosis, risk factors and disease comorbidities: A large US-based cohort study. *Frontiers in public health, 10*, 1029190.
45. Maraqa, B., Al-Shakhra, K., Alawneh, M., Jallad, R., & Alkaila, M. (2021). Demographic factors associated with COVID-19-related death in Palestine. *Public Health in Practice, 2*, 100145.
46. OSDH 2023. COVID-19 (340). COVID-19. Modified February 22 2023. Retrieved March 1, 2023.
47. Aston, A., 2024. Oklahoma tops 20,000 covid-19 deaths, 2nd highest death rate in the nation. The Oklahoman. Modified January 29, 2024. Retrieved April 18, 2024.
48. Chen, Y., Zhu, M., Zhou, Q., & Qiao, Y. (2021). Research on spatiotemporal differentiation and influence mechanism of urban resilience in China based on MGWR model. *International Journal of Environmental Research and Public Health, 18*(3), 1056.
49. Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2009). Geographically weighted regression. *The Sage handbook of spatial analysis, 1*, 243-254.
50. Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis, 28*(4), 281-298.
51. Middya, A. I., & Roy, S. (2021). Geographically varying relationships of COVID-19 mortality with different factors in India. *Scientific Reports, 11*(1), 7890.
52. Dutta, I., Basu, T., & Das, A. (2021). Spatial analysis of COVID-19 incidence and its determinants using spatial

modeling: A study on India. *Environmental Challenges, 4,* 100096.
53. Lak, A., Sharifi, A., Badr, S., Zali, A., Maher, A., Mostafavi, E., & Khalili, D. (2021). Spatio-temporal patterns of the COVID-19 pandemic, and place-based influential factors at the neighborhood scale in Tehran. *Sustainable cities and society, 72,* 103034.
54. Sheu, M. L., Hu, T. W., Keeler, T. E., Ong, M., & Sung, H. Y. (2004). The effect of a major cigarette price change on smoking behavior in California: a zero-inflated negative binomial model. *Health Economics, 13*(8), 781-791.
55. Kompaniyets, L., Pennington, A. F., Goodman, A. B., Rosenblum, H. G., Belay, B., Ko, J. Y., ... & Gundlapalli, A. V. (2021). Underlying medical conditions and severe illness among 540,667 adults hospitalized with COVID-19, March 2020–March 2021. *Preventing chronic disease, 18*, E66.
56. McConnell, K. H., Hajat, A., Sack, C., Mooney, S. J., & Khosropour, C. M. (2023). Association between any underlying health condition and COVID-19-associated hospitalization by age group, Washington State, 2020–2021: a retrospective cohort study. *BMC Infectious Diseases, 23*(1), 193.
57. Bell, T. D. (2022). COVID-19 in the critically ill patient. Infectious Disease Clinics of North America, 36(2), 365.
58. Mecenas, P., Bastos, R. T. D. R. M., Vallinoto, A. C. R., & Normando, D. (2020). Effects of temperature and humidity on the spread of COVID-19: A systematic review. PLoS one, 15(9), e0238339.
59. Wu, Y., Jing, W., Liu, J., Ma, Q., Yuan, J., Wang, Y., ... & Liu, M. (2020). Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. Science of the Total Environment, 729, 139051.
60. Luo, W., Majumder, M.S., Liu, D., Poirier, C., Mandl, K.D., Lipsitch, M. and Santillana, M., 2020. The role of absolute humidity on transmission rates of the COVID-19 outbreak. MedRxiv, 2020-02.
61. Han, J., Meyer, B. D., & Sullivan, J. X. (2020). Income and Poverty in the COVID-19 Pandemic (No. w27729). National Bureau of Economic Research.
62. Cohen-Cline, H., Li, H. F., Gill, M., Rodriguez, F., Hernandez-Boussard, T., Wolberg, H., ... & Vartanian, K. (2021). Major disparities in COVID-19 test positivity for patients with non-English preferred language even after accounting for race and social factors in the United States in 2020. BMC Public Health, 21, 1-9.
63. Choi, W. Y. (2022). Mortality rate of patients with COVID-19 based on underlying health conditions. Disaster medicine and public health preparedness, 16(6), 2480-2485.
64. Fareed, Z., Iqbal, N., Shahzad, F., Shah, S. G. M., Zulfiqar, B., Shahzad, K., ... & Shahzad, U. (2020). Co-variance nexus between COVID-19 mortality, humidity, and air quality index in Wuhan, China: new insights from partial and multiple wavelet coherence. Air Quality, Atmosphere & Health, 13, 673-682.
65. Kodera, S., Rashed, E. A., & Hirata, A. (2020). Correlation between COVID-19 morbidity and mortality rates in Japan and local population density, temperature, and absolute humidity. International journal of environmental research and public health, 17(15), 5477.
66. Lundberg, D. J., Wrigley-Field, E., Cho, A., Raquib, R., Nsoesie, E. O., Paglino, E., ... & Stokes, A. C. (2023). COVID-19 mortality by race and ethnicity in US metropolitan and nonmetropolitan areas, March 2020 to February 2022. JAMA network open, 6(5), e2311098-e2311098.
67. Jahn, B., Friedrich, S., Behnke, J., Engel, J., Garczarek, U., Münnich, R., ... & Friede, T. (2022). On the role of data, statistics and decisions in a pandemic. AStA Advances in Statistical Analysis, 106(3), 349-382.
68. Alsaqqa, H. H. (2022). Sustaining the public health intervention strategies in confronting the Covid-19 pandemic. Journal of Public Health Research, 11(2), 22799036221102493.
69. Friedrich S, Behnke J, Engel J, Garczarek U, Münnich R, Pauly M, Wilhelm A, Wolkenhauer O, Zwick M, Siebert U., 2022. On the role of data, statistics and decisions in a pandemic. AStA Advances in Statistical Analysis, 106(3):349-82.