IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

U.S. Geography & Healthcare: Infrastructure and Outcomes

Jake M. Dong

Sage Hill School, 20402 Newport Coast Dr, Newport Coast, CA 92657, USA; jakedong080218@gmail.com
Mentor: Dr. Tammy Ingram

ABSTRACT: Health disparities in the United States are often attributed to income, education, and race, but geography is a less studied dimension of inequity. This paper examines rurality on health conditions, outcomes, and infrastructure across over 3,000 U.S. counties. Using county-level data on chronic disease prevalence, mortality, life expectancy, infant mortality, and healthcare access, we ran regression models to examine the relationship between the rural population share and these health measures. Results show that rural and urban populations are equally likely to face chronic conditions such as obesity, diabetes, and high blood pressure. However, rural counties exhibit worse outcomes, with higher all-cause mortality, modestly shorter life expectancy, and especially elevated infant mortality. Preventive care also shows a rural disadvantage: cholesterol screening rates decline consistently with rurality, while other access measures, such as insurance coverage and doctor visits, remain stable. These findings indicate that rurality is linked to differences in health outcomes even when underlying health conditions appear similar. While the mechanisms behind these associations cannot be determined from the present data, the patterns highlight the need for further research into healthcare capacity, preventive services, and maternal care in rural communities.
KEYWORDS: Behavioral And Social Sciences, Sociology, Public Health, Public Policy, Geography.

Introduction

U.S. health disparities have long been a central research and political topic. Across the population, substantial differences in health outcomes are observed along social, economic, and geographic lines, contributing to persistent gaps in life expectancy, mortality, and quality of life. While overall national health indicators have improved over the past several decades, these gains have not been distributed equally. Specific populations continue to experience disproportionate burdens of disease and premature death, underscoring the importance of identifying and addressing the structural determinants of health inequity.

Robust research has demonstrated that socioeconomic factors are among the most powerful predictors of health outcomes. It has been shown that adults with lower levels of educational attainment experience higher mortality rates and shorter life expectancy across gender and racial groups.[1] Additionally, life expectancy among the wealthiest Americans exceeded that of the poorest by more than a decade, with the gap widening over time and varying substantially across regions, particularly for low-income populations.[2] There are still persistent differences in life expectancy by race and ethnicity between 2000 and 2019, with American Indian and Alaska Native populations showing slight improvements over time.[3] Similarly, there are also disproportionately higher cause-specific mortality rates among Black, American Indian, and Alaska Native populations relative to White populations, reflecting the cumulative effects of structural racism and unequal access to care.[4] This literature demonstrates that health outcomes are socially patterned, with socioeconomic status and race serving as well-established axes of inequality in the U.S.

Although socioeconomic and racial disparities are well-documented, rural-urban differences have often been treated as secondary variables rather than primary drivers of health outcomes. However, recent studies suggest that rurality may be more significant than previously assumed. It has been found that infant mortality rates are significantly higher in rural areas than in urban counties, with rural infants facing persistent disadvantages even after accounting for maternal demographics.[5] Furthermore, rural counties have consistently elevated mortality rates and shorter life expectancy than urban areas, and this gap has widened over the past several decades.[6] In a CDC guest editorial, it has also been stated that rural counties face persistent challenges such as limited access to physical and mental health care and inadequate infrastructure.[7] The above research suggests that rurality may play a bigger role in explaining health disparities than previously thought. We seek to add to this literature by examining whether rurality alone impacts health outcomes or if it compounds existing socioeconomic factors in ways that consistently disadvantage rural communities.

This paper examines rurality as a less-explored dimension of U.S. health disparities to clarify its role for policymakers. Analyzing county-level data on chronic conditions, mortality, life expectancy, infant mortality, and healthcare infrastructure asks whether geography predicts health outcomes. The results show that rural and urban residents are equally likely to face common chronic conditions, but rural counties experience measurably worse outcomes, particularly in mortality and infant survival. This suggests that the disparity arises less from underlying health status than from differences in healthcare infrastructure and the quality of care. Recognizing geography as an underexplored but policy-relevant factor can help design interventions that improve not only access, but also the effectiveness of care in rural communities.

Data

We draw the data on the rural population percentage in each mainland U.S. county from the 2020 Census Bureau on Urban vs. Rural areas in the U.S. This will allow us to directly examine the correlation between rurality and different health outcomes across the country.

CDC's PLACES database gives us data on Heart Disease (CHD), Diabetes, High Blood Pressure (HBP), Obesity, Cholesterol Screening (CS), Lack of Health Insurance (LHI), Routine Doctor Check-ups (RDC), and High Blood Pressure (HBP) Medication use. The data are presented as percentages of the adult population, and we use age-adjusted statistics for standardized comparisons. NIH's HDPulse database, updated in 2025, provides the Age-Adjusted All-Cause Mortality Rate (ACMR) per 100,000 by county. Lastly, we use the CountyHealthRankings 2025 data release for Life Expectancy (in years) and Infant Mortality Rate (IMR) per 1000 by county.

Method

We analyze data on health outcomes and ruralness by running single-variable linear regressions to examine the correlation between the percentage of the rural population and each health category. The results are correlation-only, and no inference of cause-and-effect is possible.

CHD, Diabetes, HBP, and Obesity are proxies for the health of the population in each county. These four health conditions were selected because they are not heavily influenced by genetics. For example, 90-95% of diabetes patients are Type 2, which is caused primarily by lifestyle choices. These common health conditions allow us to assess whether a rural population is healthier than an urban one, or vice versa, without accounting for genetic influences. The analysis of 3,144 counties helps smooth out any outliers or genetic fluctuations.

ACMR, Life Expectancy, and IMR can provide great insights for healthcare quality. The ACMR and Life Expectancy can give us clues about the overall quality of healthcare support one receives. IMR can have direct implications for prenatal, infant, and specialized healthcare.

Lastly, we chose CS, LHI, RDC, and HBP Medicine as representations for healthcare accessibility. CS and HBP medications are selected because they directly correlate with the health outcomes we are measuring. We also chose RDC and LHI as measurements because they are the most basic forms of healthcare.

Because of the U.S.'s many county outlining and planning systems, we imported all relevant data into a single master database and sorted the states and their respective counties alphabetically. Accordingly, we added each corresponding health data for that county in the succeeding columns. However, many rows are missing one or more data points in the database, as each database uses a different method to measure counties, municipalities, and urban planning regions. So, although there are 3144 official U.S. counties, our database contains more than 3264 rows due to the multiple differences in county criteria. For example, Baltimore County is the official census county. However, CountyHealthRankings treats Baltimore County and Baltimore City as separate entities, each with its own values, creating gaps in data when different databases are combined. We account for this by including all possible counties in each dataset and filling in the corresponding data, leaving entries blank if one database does not cover that county. These blanks are not included in our regression models, so they will not affect our results.

Results

Health Conditions:

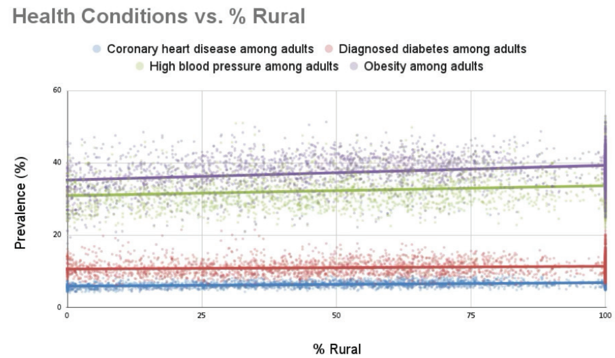

Figure 1: Scatterplots and linear regressions showing the association between county rural population share and four chronic conditions (obesity, high blood pressure, diabetes, and coronary heart disease). All correlations are weak, with small effect sizes. These results indicate that chronic disease prevalence is broadly similar across rural and urban counties, with only minimal increases as rurality rises.

Figure 1 above shows the correlations between the percentage of the rural population and the four health conditions we measured. The purple datapoints and trendline are for obesity (slope = 0.041), the blue for HBP (slope = 0.027), the red for diabetes (slope = 0.0077), and the green for CHD (slope = 0.0099). These regressions show minimal to modest urban-rural differences across chronic health conditions. For every percent increase in ruralness, there is only about a 0.01% increase in CHD, and only about 13.5% of the variations in CHD between counties are explained by ruralness (R^2 = 0.135). The trend is the same with diabetes: every percentage increase in ruralness correlates with a 0.008% increase in diabetes prevalence, and only 1.3% of the variation is explained by rurality (R^2 = 0.013). HBP rose marginally with ruralness, with each percentage increase in ruralness correlating with a 0.03% increase in prevalence, but ruralness accounted for only 3.8% of the variance (R^2 = 0.038). Obesity had the strongest correlation, with every percentage increase in ruralness associated with a 0.04% increase in obesity prevalence, though rurality explained only 8.8% of the variation (R^2 = 0.088). These results show that chronic conditions differ only slightly across rural and urban counties, and ruralness explains a small proportion of the observed variation. This suggests that, for the conditions measured, rural populations are not markedly different in underlying chronic disease prevalence, although these analyses cannot capture all dimensions of health.

DOI: 10.36838/IJHSR811.88

Health Outcomes:

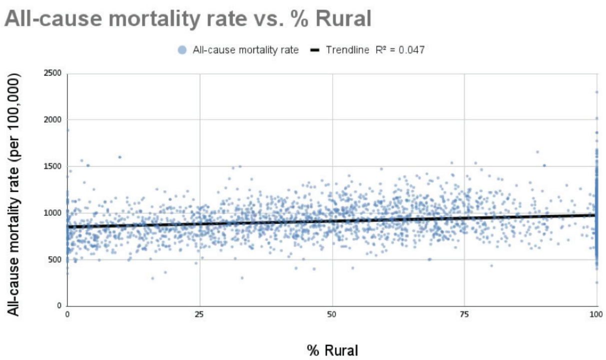

Figure 2: Scatterplot and linear regression illustrating the association between rural population percentage and age-adjusted all-cause mortality per 100,000 residents. A modest positive correlation is observed. This suggests that more rural counties experience higher mortality rates, even though the strength of the relationship remains modest.

The correlation between ACMR per 100,000 and rurality is shown in Figure 2 above. The slope of the trendline is 1.25, meaning that for every increase in rural percentage, the mortality rate increases by 1.25 per 100,000. Since typical urban-rural counties differ by 20-40% in ruralness, this corresponds to 25-50 more deaths per 100,000 people in rural counties than in urban ones. However, rurality is not the major driver of this increase in mortality rate, accounting for only 4.7% of the variation (R^2 = 0.047). These findings show a modest positive association between ruralness and mortality. When viewed alongside the weak correlations in Figure 1, this suggests that differences in mortality do not directly mirror differences in chronic condition prevalence. Because these data are correlational and lack socioeconomic and demographic controls, the underlying reasons for this pattern cannot be determined here.

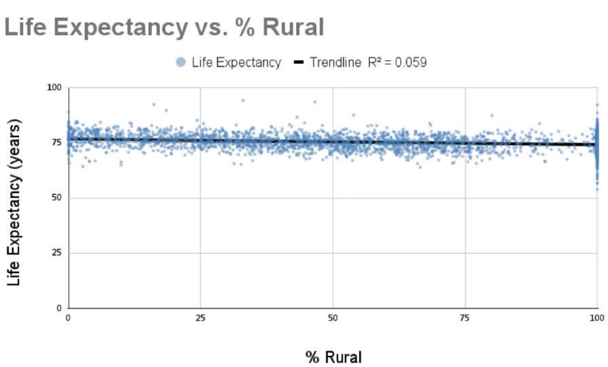

Figure 3: Scatterplot and linear regression between county rural population percentage and life expectancy at birth. The regression indicates a small negative correlation. This indicates that life expectancy tends to decline as rurality increases, though rurality explains only a small portion of the variation.

Figure 3 above shows the association between Life Expectancy (in years) and the percentage of the rural population. The trendline's slope is -0.026, meaning that each additional percentage of ruralness is associated with a decrease of around 9.5 days in life expectancy (0.026*365 ~ 9.5). Rural counties have life expectancies up to 2.6 years shorter (0.026*100 = 2.6). However, rurality alone only accounts for 5.9% of the variance (R^2 = 0.059). This finding aligns with the ACMR results in indicating that life expectancy tends to be lower in more rural counties, although the explanatory power is modest. As with other outcomes, these correlations do not indicate causation, and additional factors not included in this study likely play substantial roles.

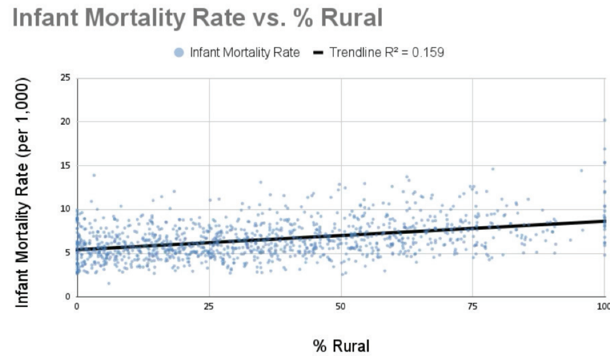

Figure 4: Scatterplot and linear regression showing the relationship between rural population percentage and infant mortality (deaths per 1,000 live births). This is the strongest correlation among the measured outcomes, though it remains modest. This highlights that infant mortality is the outcome most strongly associated with rurality, suggesting potential disparities in maternal or specialized care.

Figure 4 above shows the association between IMR per 1000 and ruralness. The slope of the trendline is 0.033, meaning that for each increase in rural population percentage, there is an additional 0.033 infant death per 1000 live births. In a county that's 40% more rural, there will be 1.3 more deaths per 1000, which is a 23% increase over the U.S. average of 5.6 per 1000. Furthermore, rurality alone accounts for 16% of the variance (R^2 = 0.16), far more than any other metric we measured. IMR shows the strongest association with rurality, matching previous research on geographic differences in infant outcomes. While this association is notable, this analysis cannot determine underlying mechanisms, and factors such as socioeconomic conditions, maternal characteristics, and healthcare access may contribute to the observed pattern.

Health Infrastructure:

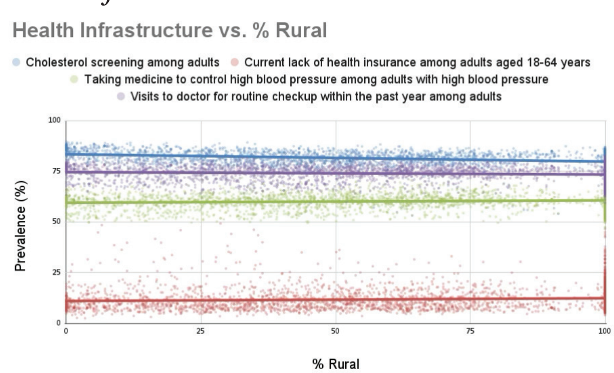

Figure 5: Scatterplot and linear regression between rural population percentage and four access measures (cholesterol screening, routine doctor checkups, lack of health insurance, and high blood-pressure medication use). Most indicators show minimal correlation, except for a modest decline in cholesterol screening. These results suggest that most basic healthcare access measures are similar across geographies, with the main difference appearing in preventive care such as cholesterol screening.

DOI: 10.36838/IJHSR811.88

Figure 5 above shows the correlation between CS, RDC, LHI, and HBP Medication and the percentage of the rural population. The blue datapoints and trendline are for CS (slope = -0.036), purple for RDC (slope = -0.012), red for LHI (slope = 0.014), and green for HBP medication use (slope = 0.012). The very flat slopes mean that the healthcare infrastructure we measured showed minimal association with rurality, with the only modest association being CS, where every percent increase in ruralness correlates to 0.036% fewer adults screened for cholesterol. This means that in a county that's 40% more rural, roughly 1.4% fewer adults will be screened. Rurality also accounts for 11.9% of the variance (R^2 = 0.119). These findings suggest that basic healthcare access indicators such as insurance coverage, doctor visits, and medication use are relatively similar across geographies. The modest association with cholesterol screening indicates possible differences in preventive care, but the dataset does not allow us to determine the causes of these differences.

Discussion

In this study, we present the correlation between four health conditions (Diabetes, Obesity, HBP, and CHD), three health outcomes (ACMR, Life Expectancy, and IMR), and four health infrastructures (CS, LHI, HBP medication, and RDC) and rurality.

Results Implications:

When we look at the chronic conditions, ruralness only shows weak correlations. The R^2 values never exceed 0.135, indicating that geography alone accounts for only a slight increase in chronic disease prevalence. In other words, the chronic conditions measured here do not differ substantially between rural and urban residents. Additionally, the breadth of healthcare infrastructure is essentially the same across urban and rural counties, with access to insurance, medication, and doctor visits very similar across geographies; the only difference is preventive services, such as cholesterol screening, where there is a noticeable decline in rural counties. However, health outcomes differ between urban and rural counties despite similar levels of health and healthcare access. ACMR rises noticeably, life expectancy declines modestly, and IMR increases most strongly with rurality. These findings show that rurality is associated with worse outcomes even when underlying chronic disease prevalence is similar, though the data cannot determine the specific causes behind these patterns.

Although structural features of the rural healthcare environment, such as longer distances to hospitals, fewer specialists, or limited facility capacity, may contribute to these differences, this interpretation cannot be confirmed from the available data. The contrast between similar prevalences of chronic illnesses and worse health outcomes suggests that qualitative differences in healthcare contexts may play a role. Still, unmeasured socioeconomic and demographic factors are also likely to be important. The elevated IMR, which may reflect gaps in specialized or preventive prenatal care, along with the decline in CS prevalence, suggests that preventive care may be less accessible in rural counties. However, this remains a correlation rather than evidence of a direct causal relationship. These preventive services are important for reducing health risks and detecting diseases early, but our county-level data cannot determine how much they contribute to outcome disparities.

However, despite the outcome differences, the low R^2 values (never exceeding 0.16) show that rurality itself explains only a small portion of the variation. This suggests that other socioeconomic factors remain more significant determinants of health disparities, consistent with previous research. Rather than acting as a primary cause, rurality appears to be linked with broader social and structural conditions that may amplify existing inequalities. Overall, the results reveal a clear pattern: Rural and urban populations are similarly healthy in terms of chronic conditions, yet rural areas face worse health outcomes, especially in mortality rates and infant survival. This divergence indicates that health disparities are shaped by a combination of geographic context and factors not captured in this analysis. Future research that includes individual-level data, richer measures of healthcare capacity, and multivariable models would help clarify the mechanisms behind these associations.

Policy Implications:

These findings have several policy implications. First, they suggest that simply expanding healthcare access, such as increasing insurance coverage or the number of available doctors, may not, on its own, close the urban-rural health disparity. Instead, policymakers may need to focus on strengthening the overall capacity and quality of rural healthcare systems, as well as improving the reach of preventive and maternal care. Efforts such as expanding telehealth services, supporting rural hospitals, and incentivizing specialized healthcare practices in rural areas could help address some of the differences observed in health outcomes, even though this study cannot determine the specific causes behind those differences. Public health programs like mobile screening clinics or subsidized preventive check-ups may also help offset geographic barriers and improve service availability.

Additionally, the results highlight the need for more consistent and standardized data collection practices. For example, Connecticut reports both counties and planning regions, which complicates comparisons across datasets, and Florida does not provide HBP or HBP medication data, which limits the completeness of county-level health analyses. Policymakers could support the development of standardized guidelines for collecting and reporting county-level health data, which would help future researchers more reliably organize, compare, and analyze health indicators across states. Clearer, more uniform data standards would strengthen the evidence base needed to better address rural-urban health disparities.

Limitations:

This study has several limitations. Firstly, our data are limited to the percentage of the rural county population and do not include potential confounding factors such as income, education, race, or gender. As a result, the associations between ruralness and health outcomes cannot be interpreted causally, since many unmeasured socioeconomic and demographic factors may play a more substantial role. Based on the available data, rurality can be viewed only as correlated with existing inequities, rather than as an independent driver of them. Secondly, our county-level data may hide disparities within counties. Because counties differ widely in size, population, and demographic composition, reducing each county to a single aggregate value may overlook important community-level variation and introduce the possibility of ecological fallacy. Lastly, the Census measure of rurality, defined only as the percentage of residents living in rural areas, does not capture structural factors such as provider availability or travel distance. For these reasons, the present analysis should be interpreted cautiously, as the findings are descriptive rather than definitive. Future research using individual-level data, multivariable models, and richer measures of healthcare capacity would help clarify the mechanisms underlying the observed geographic patterns.

Conclusion

This study shows that rural residents are not substantially less healthy than urban residents with respect to the chronic conditions measured. However, they face worse outcomes, including higher mortality, shorter life expectancy, and elevated infant mortality. These patterns may reflect differences in preventive care or healthcare environments, though the present analysis cannot determine the specific causes. The findings suggest that geography is associated with variations in outcomes even when chronic conditions appear similar, underscoring the need to better understand the factors shaping rural health. For policymakers, this means that improving rural health may require not only maintaining access but also strengthening preventive services, maternal care, and the overall capacity of local healthcare systems. Addressing these challenges will be important for reducing persistent inequities and supporting healthier, longer lives for rural communities.

Acknowledgments

I would like to give my utmost thanks to my mentor, Dr. Tammy Ingram, who was very supportive in my writing of this paper. Her guidance, feedback, and suggestions were foundational in completing this thesis and helped me grow as a scholar. I would also like to thank my friends and family, who supported me on this journey as well.

References

1. Montez JK, Hummer RA, Hayward MD. Educational attainment and adult mortality in the United States: a systematic analysis of functional form. Demography. 2012 Feb;49(1):315-36. doi: 10.1007/s13524-011-0082-8. PMID: 22246797; PMCID: PMC3290920.
2. Chetty R, Stepner M, Abraham S, Lin S, Scuderi B, Turner N, Bergeron A, Cutler D. The Association Between Income and Life Expectancy in the United States, 2001-2014. JAMA. 2016 Apr 26;315(16):1750-66. doi: 10.1001/jama.2016.4226. Erratum in: JAMA. 2017 Jan 3;317(1):90. doi: 10.1001/jama.2016.18691. PMID: 27063997; PMCID: PMC4866586.
3. GBD US Health Disparities Collaborators. Life expectancy by county, race, and ethnicity in the USA, 2000-19: a systematic analysis of health disparities. Lancet. 2022 Jul 2;400(10345):25-38. doi: 10.1016/S0140-6736(22)00876-5. Epub 2022 Jun 16. PMID: 35717994; PMCID: PMC9256789.
4. GBD US Health Disparities Collaborators. Cause-specific mortality by county, race, and ethnicity in the USA, 2000-19: a systematic analysis of health disparities. Lancet. 2023 Sep 23;402(10407):1065-1082. doi: 10.1016/S0140-6736(23)01088-7. Epub 2023 Aug 3. PMID: 37544309; PMCID: PMC10528747.
5. Ely DM, Driscoll AK. Infant Mortality in the United States, 2022: Data From the Period Linked Birth/Infant Death File. Natl Vital Stat Rep. 2024 Jul 25;(5):10.15620/cdc/157006. doi: 10.15620/cdc/157006. PMID: 39412861; PMCID: PMC11513748.
6. Singh GK, Siahpush M. Widening rural-urban disparities in life expectancy, U.S., 1969-2009. American Journal of Preventive Medicine. 2014 Feb;46(2):e19-29. DOI: 10.1016/j.amepre.2013.10.017. PMID: 24439358.
7. Matthews, K. A.; Spears, K. S.; Anderson-Lewis, C. Rural Health Disparities: Contemporary Solutions for Persistent Rural Public Health Challenges. Prev. Chronic Dis. 2025, 22, 250202. https://doi.org/10.5888/pcd22.250202

Author

Jake Dong is a senior at Sage Hill School. He found his passion in public health and policy through his work encouraging healthier lifestyles for middle school kids in his community. Jake hopes to continue participating in public health and contributing to international health policies.

DOI: 10.36838/IJHSR811.88