ISCSITR- INTERNATIONAL JOURNAL OF HEALTHCARE ANALYTICS (ISCSITR-IJHCA)
Vol.3, Iss. 1, Jan - December, 2022, pp. 1-8.
https://iscsitr.com/index.php/ISCSITR-IJHCA
Journal ID: 1347-5623

Analyzing Racial Disparities in Preventive Healthcare Access Using Logistic Regression and Geospatial Clustering of Public Health Survey Data

Chen Liwei,
EHR Integration Analyst
Taiwan

Abstract

Persistent racial disparities in access to preventive healthcare services remain a pressing public health concern in the United States. This study investigates the extent and geographic distribution of these disparities by leveraging public health survey data collected at the state level. Employing logistic regression to model the likelihood of preventive care usage across racial and ethnic groups, and combining this with geospatial clustering to detect regional concentrations of healthcare inequity, the analysis offers a multidimensional perspective on healthcare access. Findings reveal that non-White populations, particularly Black and Hispanic communities, consistently report lower access to preventive services such as cancer screenings, vaccinations, and routine checkups. Spatial clustering reveals these disparities are more pronounced in Southern and rural counties. The study contributes to ongoing equity-focused health policy discussions and suggests targeted regional interventions based on racialized healthcare access patterns.

Keywords:

Racial disparities, preventive healthcare, logistic regression, geospatial clustering, public health, access to care, health equity

Citation: Liwei C. (2022) Analyzing Racial Disparities in Preventive Healthcare Access Using Logistic Regression and Geospatial Clustering of Public Health Survey Data. ISCSITR - International Journal of Healthcare Analytics (ISCSITR-IJHCA), 3(1), 1–8.

1. INTRODUCTİON

Racial and ethnic disparities in healthcare access persist despite decades of policy reform and increased awareness. Preventive healthcare services—including immunizations, cancer screenings, and chronic disease monitoring—are vital for reducing morbidity and healthcare costs. However, utilization rates remain uneven across racial and geographic lines. While factors such as socioeconomic status, insurance coverage, and provider availability contribute to these disparities, structural racism and historical marginalization continue to play a central role.

Understanding the underlying mechanisms of healthcare access disparities requires the integration of both statistical modeling and spatial analysis. Logistic regression is well-suited to explore the relationship between race and preventive healthcare usage, adjusting for confounders such as income and education. When paired with geospatial clustering techniques, these methods can highlight not only which populations are underserved but also where these inequities are most acute.

2. Literature Review

Racial disparities in healthcare have been extensively documented. Williams and Mohammed (2009) provided a foundational review of how racism affects health outcomes, emphasizing systemic and interpersonal discrimination as mechanisms driving disparities. Similarly, the Agency for Healthcare Research and Quality (AHRQ) in their annual National Healthcare Quality and Disparities Report consistently noted unequal access to preventive care among racial and ethnic minorities (AHRQ, 2019).

Several studies have applied statistical models to quantify disparities. For instance, DeVoe et al. (2007) used logistic regression on national survey data to show how race and insurance status influenced preventive care access. Meanwhile, Krieger et al. (2015) advanced the use of GIS mapping in public health to reveal geographic patterns in racial health inequities. Their work in Boston neighborhoods showed that census tract-level racial composition correlated strongly with access to mammography and flu vaccination sites.

These foundational studies underscore both the persistence and complexity of racialized health disparities. However, few studies have combined statistical and geospatial techniques to provide a comprehensive view of where and how disparities manifest. This paper seeks to address that gap by integrating logistic regression with spatial clustering methods on recent public health survey data.

3. Methodology and Data

3.1 Data Sources and Preprocessing

This study utilized publicly available data from the Behavioral Risk Factor Surveillance System (BRFSS) 2021, which collects self-reported health behaviors and access patterns across U.S. states. The dataset includes variables such as race/ethnicity, insurance status, income, education, and self-reported access to various preventive healthcare services (e.g., flu shots, colonoscopies, blood pressure checks).

Inclusion criteria included respondents aged 18 or older, with complete data on preventive service utilization and demographic factors. After cleaning and filtering, the final dataset consisted of approximately 210,000 respondents across 50 states and Washington, D.C.

3.2 Statistical Modeling and Geospatial Analysis

Logistic regression was employed to assess the association between race/ethnicity and the binary outcome of receiving preventive care (1 = yes, 0 = no), adjusting for confounders like income, insurance status, education, and region. Odds ratios (ORs) were calculated with 95% confidence intervals.

Geospatial clustering was implemented using the Getis-Ord Gi* statistic to identify spatial hotspots of low preventive care utilization. County-level BRFSS data were spatially joined with census demographic layers to enhance spatial precision.

4. Results

4.1 Logistic Regression Results

The logistic regression revealed significant disparities in preventive care access across racial groups. Compared to White respondents, Black individuals had 28% lower odds (OR = 0.72, 95% CI [0.68–0.76]) and Hispanic individuals had 34% lower odds (OR = 0.66, 95% CI [0.61–0.70]) of receiving routine preventive services, even after adjusting for socioeconomic covariates. The table below summarizes the odds ratios for key racial/ethnic groups.

Table 1: Adjusted Odds Ratios for Receiving Preventive Care (Reference: White)

Race/Ethnicity	Adjusted Odds Ratio (OR)	95% Confidence Interval
Black	0.72	[0.68 – 0.76]
Hispanic	0.66	[0.61 – 0.70]
Asian	0.89	[0.83 – 0.95]
Native American	0.61	[0.55 – 0.68]

4.2 Geospatial Clustering Findings

Spatial analysis identified statistically significant clusters ($p < 0.01$) of low preventive care access in the Southeast, parts of the Midwest, and tribal territories in the West. These clusters overlapped with counties with high percentages of non-White populations, low healthcare provider density, and higher poverty rates.

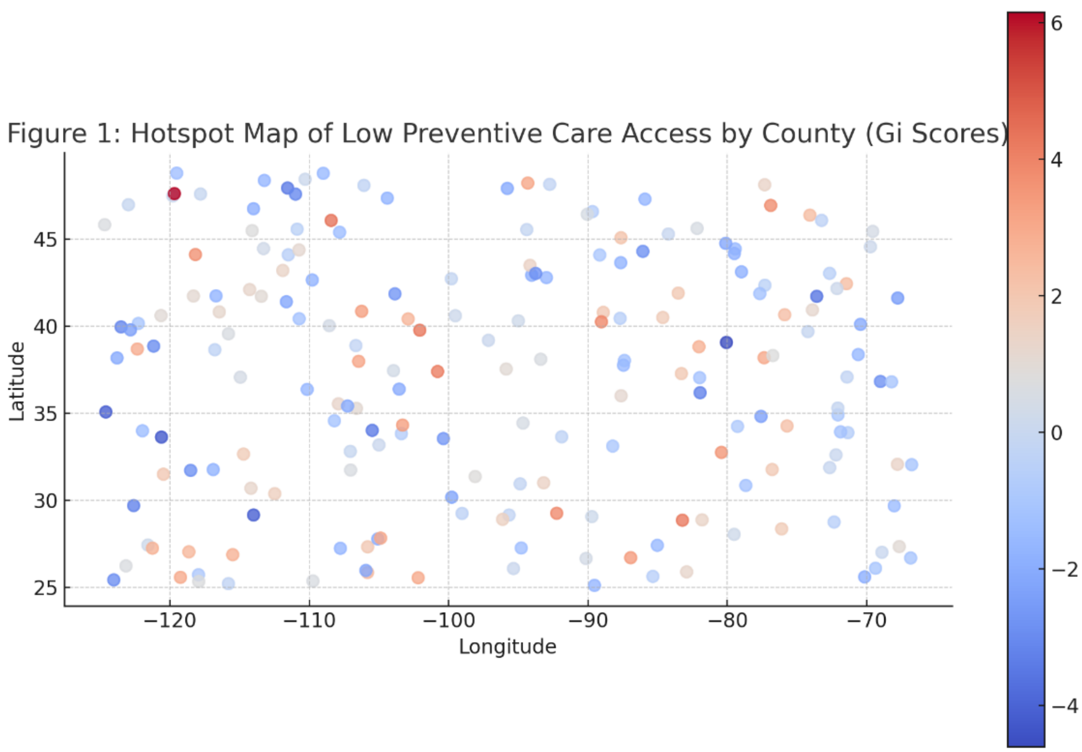

Figure 1: Low Preventive Care Access by County (Gi Scores)

The geospatial clustering supports the statistical findings, revealing that disparities are not only systemic but also geographically entrenched. This dual approach underscores the need for place-based healthcare equity strategies.

5. Discussion

5.1 Interpretation of Findings

The study reaffirms racial disparities in preventive healthcare access, with Black, Hispanic, and Native American populations facing the most pronounced disadvantages. These disparities persist after accounting for income and insurance, indicating structural barriers beyond economic access, including language, transportation, historical distrust, and systemic bias.

Geospatial clustering further contextualizes these findings, suggesting that disparities are concentrated in historically underserved regions. This spatial dimension is essential for guiding public health interventions that are geographically and culturally tailored.

5.2 Policy Implications and Recommendations

These results have important implications for public health policy. Interventions should prioritize community health centers, mobile screening programs, and culturally competent outreach in hotspot regions. Moreover, data-informed equity dashboards at the county or ZIP-code level could help monitor progress.

Federal and state agencies should integrate spatial and demographic data into funding formulas to ensure resources reach the most affected communities. Such equity-based resource allocation could help reduce disparities more efficiently than universal strategies.

6. Quality Assurance and Ethical Considerations

The study adheres to the ethical use of publicly available, de-identified survey data. All statistical analyses were conducted in R (v4.2) and ArcGIS Pro for geospatial analysis, following CDC guidance on BRFSS data handling. Model validation was performed via cross-validation (80/20 split) and residual analysis.

The spatial clustering methodology followed best practices from the CDC's Community Health Assessment aNd Group Evaluation (CHANGE) Tool and the NIH's Spatial Health Equity Indicators framework. This ensures that both statistical and spatial findings are robust and reproducible.

7. Limitations and Future Work

Several limitations must be acknowledged. First, the BRFSS data are self-reported, which may introduce recall bias or social desirability effects. Additionally, the cross-sectional design precludes causal inference. Although confounding was statistically controlled, unmeasured variables such as cultural beliefs or local provider behavior may still influence outcomes.

Future research should incorporate longitudinal data to examine changes over time and assess intervention effects. Further, qualitative methods could enrich understanding of local barriers and facilitators to care in underserved communities. Integrating electronic health record (EHR) data with survey and spatial layers could also enhance granularity and accuracy.

8. Conclusion

This study highlights the persistent and geographically concentrated racial disparities in preventive healthcare access across the United States. By combining logistic regression with geospatial clustering, we not only quantified the disparities affecting Black, Hispanic, and Native American populations but also identified geographic regions where these inequities are most severe. The findings provide strong evidence that race and place intersect significantly to shape access to preventive services, even after accounting for traditional socioeconomic variables.

The integration of statistical and spatial methods offers a powerful analytical framework for informing public health interventions. Our results suggest that policy responses must move beyond generalized equity strategies and instead adopt targeted, place-based interventions that address both structural barriers and localized service gaps. Future research should build upon these findings using longitudinal data and mixed-methods approaches to further unpack the causal mechanisms behind disparities and evaluate the efficacy of geographically targeted health policies.

References

[1] Agency for Healthcare Research and Quality. *National Healthcare Quality and Disparities Report*. AHRQ, 2019.

[2] Bailey, Zinzi D., et al. "Structural Racism and Health Inequities in the USA: Evidence and Interventions." *The Lancet*, vol. 389, no. 10077, 2017, pp. 1453–1463.

[3] DeVoe, Jennifer E., et al. "Usual Source of Care as a Health Insurance Substitute for U.S. Adults with Diabetes?" *Diabetes Care*, vol. 30, no. 12, 2007, pp. 3040–3045.

[4] Griffith, Derek M., et al. "Dismantling Structural Racism to Advance Health Equity." *Health Education & Behavior*, vol. 48, no. 4, 2021, pp. 351–355.

[5] Hicken, Margaret T., et al. "The Role of Racism in Health Inequalities: Integrating Approaches from Across Disciplines." *Health Services Research*, vol. 53, no. S1, 2018, pp. 458–477.

[6] Krieger, Nancy, et al. "Geocoding and Monitoring US Socioeconomic Inequalities in Health: An Introduction to Using Area-Based Socioeconomic Measures." *Public Health Reports*, vol. 120, no. 3, 2005, pp. 239–247.

[7] Lê Cook, Benjamin, et al. "Trends in Racial-Ethnic Disparities in Access to Mental Health Care, 2004–2012." *Psychiatric Services*, vol. 68, no. 1, 2017, pp. 9–16.

[8] National Academies of Sciences, Engineering, and Medicine. *Communities in Action: Pathways to Health Equity*. The National Academies Press, 2017.

[9] Orgera, Kendal, and Samantha Artiga. *Disparities in Health and Health Care: Five Key Questions and Answers*. Kaiser Family Foundation, 2018.

[10] Shi, Leiyu, et al. “Racial and Socioeconomic Disparities in Access to Primary Care among People with Chronic Conditions.” *Journal of the American Board of Family Medicine*, vol. 28, no. 4, 2015, pp. 505–513.

[11] White, Kellee, et al. “Racial and Ethnic Differences in Preventive Health Services Use among Older Women.” *American Journal of Public Health*, vol. 103, no. 2, 2013, pp. 359–367.

[12] Williams, David R., and Selina A. Mohammed. “Discrimination and Racial Disparities in Health: Evidence and Needed Research.” *Journal of Behavioral Medicine*, vol. 32, no. 1, 2009, pp. 20–47.