



# The Positive Relationship of Walkability on Diabetes Prevalence in the Southern United States

A. Bose, S. Oberg, A. Cheruvu (Department of Statistics)

## I. ABSTRACT

The diabetes epidemic in the United States presents a nuanced public health challenge, shaped by factors such as socioeconomic status and climate. While the influence of these factors on diabetes is well-established, the role of walkability in managing diabetes prevalence remains contested. This study revisits the relationship between walkability and diabetes in the U.S., using walkability indexes calculated from CDC data. Contrary to some studies suggesting that increased walkability reduces diabetes prevalence, our findings, analyzed through Geographically Weighted Regression (GWR), reveal that walkability is not a significant predictor of diabetes prevalence and exhibits notable regional anomalies. Further analysis using Monte Carlo simulations, Global I Moran’s Test, and Variance Inflation Ratio (VIR) supports these results. Our study also critiques the current methods of calculating the walkability index, proposing a revised model that incorporates additional relevant variables from the CDC. This nuanced understanding underscores the need for region-specific urban planning and public health strategies that recognize the complex interplay between walkability, environmental, and socioeconomic factors.

## II. INTRODUCTION

Diabetes, a chronic illness linked to high blood sugar levels, affects approximately 9.7% of U.S. adults, impacting both physical and mental well-being and straining healthcare resources. Recent studies suggest that the walkability of a region—a measure influenced by factors like intersection density and transit proximity—might significantly impact diabetes prevalence. While some findings indicate that higher walkability scores are associated with lower diabetes rates, our study explores this relationship in the Southern United States. Contrary to earlier research, we found that increased walkability correlates with higher diabetes prevalence in this region, suggesting that geographic and socioeconomic factors may influence health outcomes differently across locations. Understanding these dynamics is essential for devising effective public health strategies to manage and prevent diabetes.

## III. METHODS

### i. Data Sources

Data for this study were gathered from multiple sources. The Walkability Index Data were obtained from the Environmental Protection Agency (EPA), utilizing the current formula provided by the agency. Diabetes prevalence data were sourced from the Centers for Disease Control and Prevention (CDC), while temperature data were acquired from Geographic Information Systems (GIS) for Racial Equity. Median household income information was retrieved from the Small Area Income and Poverty Estimates provided by the United States Census Bureau.

### ii. Statistical Analysis

To explore the spatial relationships between walkability index, diabetes prevalence, temperature, and median household income, we employed Geographic Weighted Regression (GWR). GWR allows for the examination of spatially varying relationships between variables. Additionally, to assess the robustness of our GWR results, Monte Carlo simulation was utilized.

### iii. Spatial Diagnostics

Global Moran’s I test was conducted to examine spatial autocorrelation in the variables, providing insights into the spatial patterns present in the data. Furthermore, Variance Inflation Factor (VIF) analysis was employed to diagnose multicollinearity among the predictor variables.

### iv. Enhancing Walkability Measures

In identifying alternate factors that could be used to calculate the walkability index, we considered the limitations of the current formula. Factors such as sidewalk conditions, access to public transportation, crime rates,

and availability of green spaces were identified as potential contributors to walkability. Integrating these factors into the walkability index could enhance its accuracy and applicability, providing a more comprehensive understanding of the factors influencing walkability in the studied areas.

## IV. RESULTS

Our study applied a Geographically Weighted Regression (GWR) model to explore the connection between walkability index score and diabetes prevalence across the U.S. Simulations with artificial data validated the model’s effectiveness, showing accurate, reliable predictions with evenly dispersed residuals. Analysis of real-world data highlighted a notable positive correlation between walkability and diabetes prevalence in the Southern United States, where higher walkability is associated with increased diabetes rates, as depicted in **Figure 1**. This contrasts with other regions like the West Coast and Pacific Northwest, where higher walkability tends to decrease diabetes prevalence. Ultimately, we believed that there would be a consistent trend showing that the Walkability Index Score did impact diabetes across the entire U.S. However, from our GWR model’s results, that that is not the case. This leads us to believe that the Walkability Index Score calculated by the CDC may not be truly accurate and may fail to account for other additional factors that could contribute to increased diabetes prevalence in the U.S.

### i. Validation

This is further shown when conducting a Monte-Carlo test. As seen in **Figure 1**, the National Walkability Index Score is associated with a p-value of 0.18. This value indicates that the National Walkability Index Score is not a statistically significant metric as its p-value is greater than 0.05. Thus, the Walkability Index Score may not have a significant impact on predicting diabetes prevalence across the U.S. In addition, from the Global Moran-Eye Test, our model’s Moran 1 Statistic is 0.049, meaning there is slight positive autocorrelation in the covariates. However, we expect this as our covariates tend to resemble similar factors(eg.smoking,obesity,high blood pressure, etc.). However, this is not something to be too worried about as there is not significant autocorrelation or high spatial dependence. In addition, other validation metrics we looked at include the Variance Inflation Factor(VIF) of each of our covariates. None of our covariates had a VIF score that exceeded 13, indicating there is not significant multicollinearity in our GWR model that poses an issue. We also looked at the residual plot and we can see that there are fairly evenly dispersed residuals which is a positive sign.

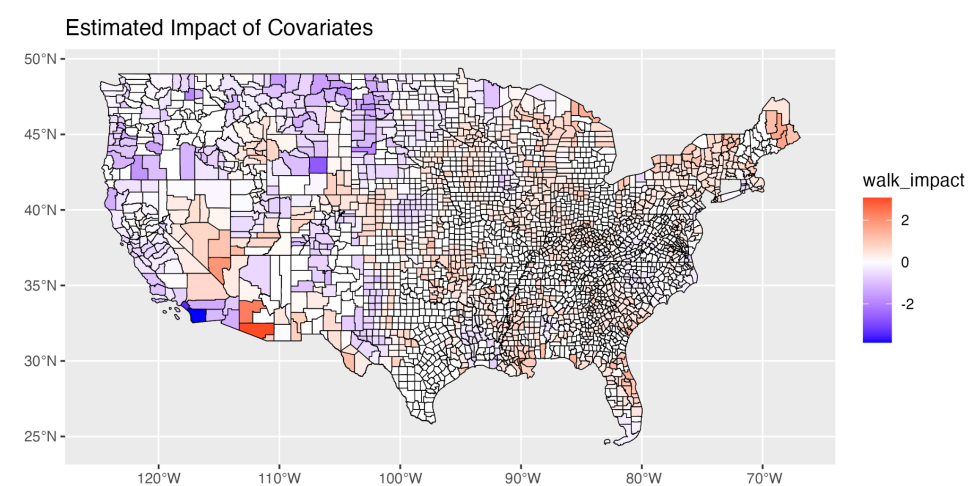


Figure 1: Figure 3

Regional factors such as climate play a significant role; in warmer Southern climates, less outdoor activity may lead to higher diabetes rates, while cooler climates show the opposite effect. Additional analyses identified other influencing factors, including obesity and smoking, which significantly impact diabetes rates, as shown in **Figure 2**.

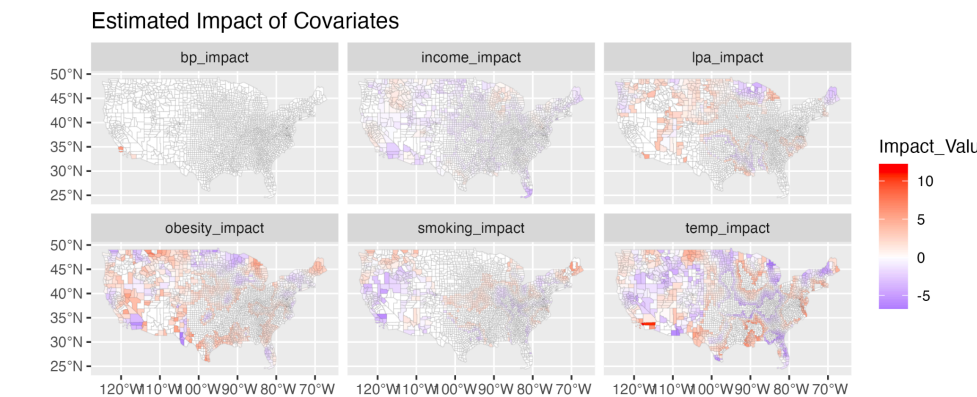


Figure 2: Impact of Covariates on Diabetes Prevalence

These insights underscore the complex relationship between environmental factors and health outcomes, highlighting the necessity for region-specific public health strategies.

## V. DISCUSSION

Our investigation into the relationship between walkability and diabetes prevalence in the U.S. revealed that walkability, as currently indexed, does not significantly predict diabetes outcomes and varies notably across different regions. This challenges the dominant assumption that higher walkability universally lowers diabetes risk and aligns with our analytical results from Geographically Weighted Regression (GWR), Monte Carlo simulations, Global I Moran’s Test, and Variance Inflation Ratio (VIR) assessments.

### i. Critique of Current Walkability Indices

Our findings question the efficacy of current walkability indices used by public health and urban planning bodies, suggesting that these indices fail to capture the multifaceted influences on health outcomes effectively. We propose a revised model for calculating walkability that incorporates a broader range of variables provided by the CDC, which could more accurately reflect the true impact of walkability on health.

### ii. Tailoring Public Health Strategies

Given the nuanced relationship between walkability and diabetes prevalence discovered in our research, there is a need for public health strategies that are tailored to local conditions. Urban planning should not only promote walkability but also ensure that the quality of walkable areas is good for to promoting health, integrating elements such as green spaces to enhance lifestyle quality.

### iii. Necessity for Region-Specific Approaches

Our research highlights the necessity for developing public health policies and urban planning initiatives that respect regional specificities. The variability in how walkability affects diabetes prevalence across different regions highlights the inadequacy of a one-size-fits-all approach and advocates for interventions that are informed by local data to ensure that strategies are both relevant and impactful.