Retail Food Environment and County-Level Cardiac Mortality in the United States: A Nationwide Ecological Study

Article type: Original Article

Author names:

1. Jeevan Nammi
2. Roshini Pasala
3. Ausrit Datta Poruri
4. Ramakanth Vasam
5. Nikhil Andhe

Degrees and Affiliations:

1. MBBS, University of Alabama at Birmingham, Birmingham, United States
2. MBBS, University of Alabama at Birmingham, Birmingham, United States
3. MBBS, Siddhartha Medical College, Vijayawada, India
4. MBBS, Siddhartha Medical College, Vijayawada, India
5. MBBS, Siddhartha Medical College, Vijayawada, India

ORCID (Open Researcher and Contributor Identifier):

1. https://orcid.org/0000-0003-1936-4210
2. https://orcid.org/0009-0009-4832-362X
3. https://orcid.org/0009-0009-1391-9714
4. https://orcid.org/0009-0002-1592-4495
5. N/A

About the author: Jeevan Nammi, MBBS, is a postdoctoral research fellow in Endocrinology at the University of Alabama at Birmingham (UAB), where he is also pursuing a Master of Public Health in Population Health. An ECFMG-certified international medical graduate, he has received multiple research recognitions, including the ENDO Young Investigator Award, ENDO Early Career Award and currently serves as Chair of the UAB Postdoctoral Association.

Corresponding author: Jeevan Nammi, MBBS, University of Alabama at Birmingham, Birmingham, United States. Email: jeevannammi1010@gmail.com

Acknowledgment: None

Funding: This study received no specific external funding and did not receive NIH or NIHR grant support. None of the authors are employed by the U.S. or U.K. government.

Conflict of interest statement by authors: The authors declare no financial or non-financial conflicts of interest relevant to the research, authorship, or publication of this manuscript.

Authors Contribution Statement: Conceptualization: JN. Data Curation: JN, RP, AD, RV, NA. Formal Analysis: JN, RP. Investigation: JN, RP. Methodology: JN. Project Administration: JN, RP. Resources: JN, RP. Software: JN, RP. Supervision: JN, RP. Validation: JN, RP. Visualization: JN. Writing - Original Draft: JN, RP. Writing – Review & Editing: JN, RP.

Manuscript word count (main text): 4,771

Abstract word count: 235

Number of Figures and Tables: 3 figures, 6 main tables, and 3 supplementary tables.

Abstract

Background

Cardiovascular disease remains the leading cause of death in the United States. We examined whether county-level retail food environments were associated with cardiac mortality.

Methods

A cross-sectional ecological analysis included 2,793 U.S. counties. Age-adjusted cardiac mortality rates for 2018–2020 were obtained from CDC WONDER, and retail food environment measures were obtained from the USDA Food Environment Atlas. Univariable models used raw Retail Food Environment Index (RFEI) values. Multivariable models used z-standardized continuous predictors and HC3 heteroskedasticity-consistent standard errors. Sensitivity analyses included models excluding potential mediators, log-transformed models, and a spatial error model among 2,793 counties linked to county boundary data.

Results

The mean cardiac mortality rate was 246.98 deaths per 100,000 population (SD = 58.46). In univariable analysis, higher RFEI was associated with higher cardiac mortality (β = 1.73 deaths per 100,000 per 1-unit higher RFEI; 95% CI, 1.24–2.23). In the full conditional OLS model, a 1-SD higher RFEI was associated with 4.58 additional deaths per 100,000 population (95% CI, 2.99–6.18). Residual spatial dependence was present after OLS modeling (Moran test $\chi^2 = 368.27$; $p < 0.001$). In the spatial-error model, the association was attenuated but remained positive (β = 2.65 deaths per 100,000; 95% CI, 1.11–4.19; p = 0.001).

Conclusion

Less favorable county retail food environments were associated with higher cardiac mortality after covariate adjustment and spatial sensitivity analysis. These ecological findings identify a structural correlate rather than an individual-level or causal effect.

Keywords: Cardiovascular Diseases; Mortality; Food Supply; Food Insecurity; Cross-Sectional Studies; Socioeconomic Factors; Residence Characteristics; United States. (Source: MeSH-NLM).

Introduction

Cardiovascular disease remains the leading cause of mortality among men, women, and most racial and ethnic groups in the United States.[1] In 2022 alone, cardiovascular disease accounted for 941,652 deaths nationwide, and on average one person died from a cardiovascular event every 34 seconds, according to the 2025 statistical update.[2] These figures underscore the substantial and persistent burden of cardiovascular disease in the United States. Although most U.S. households were food secure in 2023, disparities in food quality, affordability, and accessibility remain substantial across communities.[3–5]

Access to nutritious, affordable food is a critical determinant of cardiovascular health. However, many counties across the United States face significant barriers due to disparities in the local food environment.[5-7] Two key concepts, food deserts and food swamps, highlight these structural challenges.[8-11] Food deserts are generally defined as low-income areas with limited access to supermarkets or other retailers offering fresh produce and nutrient-dense foods, whereas food swamps refer to areas in which outlets selling calorie-dense, nutrient-poor foods, such as fast-food restaurants and convenience stores, substantially outnumber healthier alternatives.[12-14] Prior research has linked such environments to population-level differences in diet quality, obesity, diabetes, and other cardiometabolic outcomes. However, these associations should be interpreted cautiously, as they reflect contextual and structural patterns rather than direct causal effects at the individual level. The American Heart Association's 2025 statistical update reported dietary cardiovascular-health scores ranging from 31.4 to 53.1 across adult demographic groups on a 0-100 scale, based on National Health and Nutrition Examination Survey data from 2013-2018. These findings highlight persistent deficiencies in dietary health, although retail outlet composition does not directly measure individual dietary intake. [2]

One commonly used measure of the local food environment is the Retail Food Environment Index (RFEI), which estimates the ratio of food outlets conventionally classified as less healthy to those considered healthier within a defined geographic area.[12-14] Higher RFEI values indicate environments more heavily dominated by

unhealthy outlets. Although RFEI is useful for large-scale ecological analyses, it remains a crude proxy for food environment quality because it does not capture the nutritional content, affordability, cultural relevance, or purchasing patterns associated with individual outlets. In addition, RFEI may be intertwined with broader structural determinants, including socioeconomic disadvantage, racialized residential patterns, transportation barriers, and healthcare access, all of which may also influence cardiovascular outcomes.

Prior studies have linked local food environments to diet quality, obesity, diabetes, cardiovascular risk factors, and mortality. Food deserts, food swamps, food insecurity, and food outlet composition describe related but distinct dimensions of the food environment. Existing studies suggest that communities with limited access to healthier foods or greater exposure to less healthy outlets may experience worse cardiometabolic outcomes; however, the literature remains heterogeneous because studies differ in geographic scale, food-environment definitions, outcome definitions, and adjustment strategies.[4-18]

At the national county level, Parekh et al. demonstrated that broader food-environment disadvantage was associated with cardiovascular mortality in the United States. Their analysis used a Food Environment Index derived from food insecurity and food desert measures and examined 2017-2019 age-adjusted cardiovascular mortality among adults aged 45 years and older across 2,615 U.S. counties[19]. This prior work established an important relationship between overall food-environment quality and cardiovascular mortality. However, because that exposure was a composite measure, it does not distinguish household food insecurity, geographic food access, and retail outlet composition as separate but overlapping dimensions of the food environment.

This gap is important because retail outlet composition is a directly measurable feature of the commercial food landscape. RFEI specifically captures the relative balance of outlets conventionally classified as less healthy, such as limited-service restaurants and convenience stores, compared with outlets classified as healthier, such as grocery stores and related food retail sources.[12-14] Evaluating RFEI may therefore clarify whether outlet composition provides additional ecological information beyond broader food-environment disadvantage. In addition, the classification of superstores remains uncertain because these outlets may provide both affordable staple foods and highly processed foods; therefore, alternative RFEI specifications may yield different inferences.

The present study was well-suited to address this gap because it examined 2018-2020 age-adjusted cardiac mortality across 2,793 U.S. counties, explicitly defined RFEI-based outlet-composition measures, evaluated alternative superstore classifications, used sequential adjustment models, and incorporated spatial-error sensitivity analysis to account for residual geographic dependence. We therefore examined whether county-level RFEI, as a measure of retail food outlet composition, was associated with age-adjusted cardiac mortality across U.S. counties, and whether this association persisted across alternative RFEI specifications, sequential adjustment models, and spatial-error sensitivity analysis.

Methods

Study Design and Data Sources

We conducted a cross-sectional ecological study at the U.S. county level, integrating multiple publicly available national datasets.[20] The primary outcome, age-adjusted cardiac mortality rates (per 100,000 population), was obtained from the CDC WONDER database for 2018-2020 and aggregated over three years to improve stability of county-level rates.[21,22] Mortality was defined using heart disease–related deaths based on ICD-10 codes (I00-I09, I11, I13, I20–I51) and age-adjusted to the 2000 U.S. standard population to allow comparability across counties.

Food environment measures, including the Retail Food Environment Index (RFEI), were obtained from the USDA Food Environment Atlas 2020 edition.[23] The underlying retail-outlet indicators represent 2016 measurements for grocery stores, convenience stores, fast-food restaurants, specialized food stores, and supercenters and warehouse club stores, and 2018 measurements for farmers markets. The 2020 designation refers to the Atlas release rather than a common measurement year. Source variables, original workbook fields, and analytic covariates are documented in **Supplementary Table 1**.

Socioeconomic, demographic, behavioral, and food access variables were also sourced from publicly available county-level files and linked using Federal Information Processing Standards (FIPS). An Open Science Framework (OSF) project was created on May 16, 2025, to document the study's proposed research (https://osf.io/jxawf/overview?view_only=b97b0d7be07e4a0e91bfe74f57a18ef3). The OSF activity record confirms that a project file was uploaded on the same date. The project description outlined the ecological study design, RFEI-based exposures, proposed health outcomes, covariates, regression analyses, and sensitivity analyses. However, the project was not submitted through the formal OSF registration workflow; therefore, the study is not described as formally preregistered.

The initial plan specified an ecological cross-sectional county-level design, RFEI-based exposure definitions, cardiac mortality as the primary outcome, adjustment for socioeconomic, demographic, behavioral, and cardiometabolic covariates, and regression diagnostics including residual analysis, variance inflation factors, and model-fit assessment. Spatial-error modeling, expanded spatial diagnostics, and additional sensitivity reporting were added during revision in response to editorial methodological concerns. Additional differences between the initial analysis plan and the final study included the operational definition of cardiac mortality, the primary RFEI denominator, and the omission of proposed healthcare-access adjustments and interaction analyses.

All analyses followed the STROBE guidelines for observational studies.[24]

Outcome Variable

The primary outcome was age-adjusted cardiac mortality for 2018-2020, defined as deaths attributable to heart disease per 100,000 population. The outcome was analysed as a continuous county-level rate.

Exposure Variables

The primary exposure was the Retail Food Environment Index (RFEI), defined as a county-level ratio of outlets conventionally classified as less healthy to those conventionally classified as healthier. The primary RFEI numerator included limited-service restaurants and convenience stores. The primary denominator included grocery stores.

For alternative exposure definitions, limited-service restaurants and convenience stores were classified as less healthy outlets, while grocery stores, farmers markets, and specialty food stores were classified as healthier outlets. These classifications are proxy measures based on outlet type and do not measure in-store product quality, pricing, marketing, or purchasing behavior. Superstores were modelled separately because they may contain both healthier food options and highly processed food products. [25]

Two alternate indices were constructed to test sensitivity to superstore classification:

- RFEI1: Superstores classified as less healthy
- RFEI2: Superstores classified as healthier outlets

In all indices, higher values indicate a greater relative concentration of outlets conventionally classified as less healthy compared with healthier food outlets. Counties with denominator values of zero were treated as having undefined index values and were excluded from analyses involving that index.

Calculation of Food indices

RFEI = (fast food restaurants + convenience stores) ÷ grocery stores

RFEI1 = (fast food restaurants + convenience stores + superstores) ÷ (grocery stores + farmers markets + specialized stores)

RFEI2 = (fast food restaurants + convenience stores) ÷ (grocery stores + farmers markets + specialized stores + superstores)

Covariates

Covariates were selected a priori based on a conceptual framework linking structural food environments, socioeconomic conditions, behavioral risk factors, and cardiovascular outcomes.

- Socioeconomic variables included median household income, poverty rate, Supplemental Nutrition Assistance Program (SNAP) participation (%), and food insecurity (%).
- Demographic variables included racial/ethnic composition (% White, % Black, % Hispanic, % Asian, % Native American, % Pacific Islander). Race and ethnicity measures were derived from the USDA Food Environment Atlas using 2010 U.S. Census data. Percentage White and percentage Black represented non-Hispanic populations, whereas percentage Hispanic represented individuals of Hispanic origin. The remaining racial categories were reported according to their USDA definitions, which did not explicitly specify non-Hispanic status. Therefore, complete mutual exclusivity across all racial and ethnic categories was not assumed.
- Health/behavioral variables included smoking prevalence (%), physical inactivity (%), diabetes prevalence (%), and obesity prevalence (%).
- Food access variables included the percentage of the population with low access to grocery stores.

Univariable models used raw RFEI values and are interpreted per 1-unit higher RFEI. Multivariable OLS models and spatial-error models used z-standardized continuous predictors; therefore, adjusted coefficients represent differences in cardiac mortality per 1-SD higher predictor value. Log-transformed models are interpreted as percent differences in cardiac mortality per 1-SD higher exposure.

We recognize that several variables, including diabetes prevalence, obesity, and physical inactivity, may lie along potential causal pathways between food environment exposures and cardiovascular outcomes. Accordingly, multivariable models are interpreted as estimating conditional associations rather than total effects. To address compositional multicollinearity among racial and ethnic percentage variables, percentage White was omitted from the fully adjusted model and treated as the reference category. Variance inflation factors (VIF) were examined to assess multicollinearity.

Data Cleaning and Exclusions

Of the 3,141 U.S. counties initially assessed, 227 were excluded because cardiac mortality was missing or suppressed, 35 additional counties were excluded because one or more RFEI indices were missing or undefined, and 86 additional counties were excluded because at least one model covariate was missing. The final complete-case analytic sample included 2,793 counties.

Data linkage across datasets was performed using county-level Federal Information Processing Standards (FIPS) codes. Complete-case analysis was used. Included and excluded counties were compared descriptively using variables available in both groups. Because missingness varied across excluded counties, available denominators were reported separately for each variable. These comparisons were used to assess potential selection differences associated with complete-case analysis (**Supplementary Table 2**).

The inclusion process is summarized in a study-unit flow diagram (**Figure 1**).

Statistical Analysis

Analyses were performed using Stata. Descriptive statistics summarized all variables. Pearson correlation coefficients were used to assess bivariate associations between food-environment indices and cardiac mortality. Univariable linear regression models used raw RFEI values. Multivariable models used z-standardized continuous predictors.

Sequential multivariable models were fitted: Model 1 included standardized RFEI only; Model 2 added socioeconomic, demographic, food-access, and smoking variables; and Model 3 added diabetes, obesity, and physical inactivity. Because diabetes, obesity, and physical inactivity may lie along causal pathways, Model 3 was interpreted as a full conditional association model rather than a total-effect model.

Model diagnostics included adjusted R-squared, variance inflation factors (VIFs), and the Breusch-Pagan/Cook-Weisberg test for heteroskedasticity. HC3 Heteroskedasticity-consistent standard errors were used for inference in the fully adjusted OLS model. The log-transformed outcome models used Stata's vce (robust) estimator. Residual spatial dependence was assessed using a Moran test on residuals from the fully adjusted OLS model. County polygon data were linked to the analytic dataset using 5-digit FIPS codes.

One shapefile record used the older Shannon County, South Dakota designation and was harmonized to the current Oglala Lakota County FIPS code before linkage. A row-normalized spatial weights matrix was constructed using queen contiguity, under which counties sharing a boundary or vertex were considered neighbors. The final spatial analytic sample included the same 2,793 counties as the primary complete-case analysis. Twelve counties had no contiguous neighbors under the contiguity definition. These counties were retained in the spatial analytic sample and assigned zero neighbor weights in the row-normalized contiguity matrix; therefore, they contributed to the non-spatial regression component but not to spatially lagged neighbor information. A spatial-error model was fitted using generalized spatial two-stage least squares with the same row-normalized county-contiguity weights matrix. A formal post-estimation Moran diagnostic for the generalized spatial two-stage least-squares spatial-error model was not available through the Stata workflow used. Consequently, residual spatial autocorrelation after spatial-error estimation was not formally assessed. The fitted spatial-error parameter and attenuation of the RFEI coefficient were not interpreted as evidence that residual spatial autocorrelation had been eliminated. In log-transformed sensitivity analyses, the natural logarithm of cardiac mortality was modeled using OLS regression with standardized predictors. These models used Stata's vce(robust) variance estimator. Regression coefficients were converted into percentage differences using $100 \times [\exp(\beta) - 1]$, with delta-method confidence intervals calculated using nlcom.

Ethical Approval

This study used publicly available, de-identified, county-level data and did not involve human subjects. In accordance with 45 CFR 46.102, the activity was determined not to require Institutional Review Board (IRB) review.

Reporting Guidelines

The study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cross-sectional studies.

Results

Data Completeness and Descriptive Statistics

The initial dataset included data from 3141 counties across the United States. After excluding counties with missing data for key variables (RFEI indices, cardiac mortality, and covariates), the final analytic sample consisted of 2793 counties. The mean age-adjusted cardiac mortality rate was 246.98 deaths per 100,000 population (SD = 58.46). The mean RFEI was 7.01 (SD = 4.34) (**Table 1**). In available-case comparisons, excluded counties had a lower mean RFEI than included counties (4.03 ± 3.30; n = 266 versus 7.01 ± 4.34; n = 2,793) and higher mean low grocery-store access (47.61% ± 35.76%; n = 329 versus 20.19% ± 14.13%; n = 2,793). Because these variables were unavailable for some excluded counties, the comparisons do not fully characterize all 348 excluded counties and cannot rule out selection bias (**Supplementary Table 2**).

Correlation Analysis

Pearson correlation analysis demonstrated a weak but statistically significant positive correlation between cardiac mortality and RFEI ($r = 0.13$, $p < 0.001$). Stronger correlations were observed for RFEI1 ($r = 0.25$, $p < 0.001$) and RFEI2 ($r = 0.24$, $p < 0.001$) (**Table 2**).

Univariable Regression

Univariable linear regression models demonstrated statistically significant positive associations between all food environment indices and cardiac mortality. For each one-unit increase in RFEI, cardiac mortality increased by 1.73 deaths per 100,000 population (95% CI 1.24-2.23; $p < 0.001$; $R^2 = 0.017$). RFEI1 and RFEI2 were also positively associated with cardiac mortality (RFEI1 $\beta = 5.33$; 95% CI, 4.55-6.10; $R^2 = 0.061$; RFEI2 $\beta = 6.38$; 95% CI, 5.44-7.33; $R^2 = 0.059$; both $p < 0.001$) (**Table 3).**

The variance explained by these unadjusted models was modest, particularly for the primary RFEI model, indicating that retail food environment characteristics represent one of several contributors to county-level variation in cardiac mortality.

Multivariable Regression

In standardized sequential models, a 1-SD higher RFEI remained associated with higher cardiac mortality. The association attenuated from 7.51 deaths per 100,000 in the RFEI-only model (95% CI, 5.32-9.71; $p < 0.001$) to 4.84 deaths per 100,000 after adjustment for structural/contextual covariates (95% CI, 3.27-6.40; $p < 0.001$) and 4.58 deaths per 100,000 in the full conditional model (95% CI, 2.99-6.18; $p < 0.001$) (**Table 4**).

In the full conditional model, higher poverty, SNAP participation, food insecurity, smoking prevalence, diabetes prevalence, and physical inactivity were associated with higher cardiac mortality. Higher median income was associated with lower mortality. Grocery access was not independently associated after adjustment.

After omitting percentage White because of the compositional nature of county racial/ethnic percentages, the maximum VIF was 5.44 for poverty rate, the mean VIF was 2.31, and the VIF for RFEI was 1.09. Individual demographic coefficients were interpreted cautiously because demographic and socioeconomic county-level variables remain structurally correlated. Pearson correlations among the predictors included in the fully adjusted model are presented in **Supplementary Table 3**.

Heteroskedasticity was present based on the Breusch-Pagan/Cook-Weisberg test (chi-square = 186.97; $p < 0.001$). Therefore, heteroskedasticity-robust standard errors were used for OLS inference.

Sensitivity Analysis

When variables considered potential mediators (diabetes, obesity, and physical inactivity) were excluded from the model, a 1-SD higher RFEI remained associated with higher cardiac mortality ($\beta = 4.84$; 95% CI, 3.27-6.40; $p < 0.001$) (**Table 5)** (**Figure 2**).

Log-transformed Models

In the log-transformed models, a 1-SD higher RFEI was associated with a 2.08% higher cardiac mortality rate (95% CI, 1.44%-2.72%; $p < 0.001$). RFEI1 and RFEI2 showed similar associations of 2.89% (95% CI, 2.15%-3.62%) and 2.63% (95% CI, 1.90%-3.36%), respectively (**Table 5**).

Model-Fit Analysis

The fully adjusted OLS model yielded $R^2 = 0.502627$ and adjusted $R^2 = 0.499941$, explaining approximately half of the observed county-level variation in cardiac mortality. The remaining variation may reflect unmeasured factors, measurement error, and other sources of heterogeneity that cannot be separately quantified in this analysis (**Table 6**).

Spatial Autocorrelation

Residual spatial dependence was present after the fully adjusted OLS model (Moran test $\chi^2 = 368.27$; $p < 0.001$). Therefore, a spatial-error model was fitted. The 12 counties without contiguous neighbors were retained in the spatial analytic sample with zero neighbor weights. In this model, a 1-SD higher RFEI remained positively associated with cardiac mortality, although the coefficient was attenuated compared with OLS ($\beta = 2.65$ deaths per 100,000; 95% CI, 1.11-4.19; $p = 0.001$). The fitted spatial-error parameter was $\rho = 0.47$ (95% CI, 0.43-0.51; $p < 0.001$). This parameter characterizes spatial dependence in the fitted model but does not establish whether residual spatial autocorrelation was eliminated. A formal Moran diagnostic on residuals from the generalized spatial two-stage least-squares spatial-error model was not available in the analytic workflow.

(**Table 6**).

Geographic Distribution

Choropleth maps demonstrate regional variation in both RFEI and cardiac mortality (**Figure 3).** These should be interpreted as descriptive visualizations of geographic variation rather than evidence of causal geographic effects.

Discussion

This nationwide county-level ecological study found that counties with higher RFEI values had higher age-adjusted cardiac mortality. The association was statistically significant in univariable analysis, persisted after adjustment for county-level socioeconomic, demographic, food-access, behavioral, and cardiometabolic covariates, and remained positive after excluding potential mediators, log-transformed cardiac mortality models, and accounting for spatially structured residual dependence. However, the magnitude of association was modest, supporting interpretation of RFEI as a county-level structural correlate of cardiac mortality rather than evidence of individual-level risk or causality.

These findings are concordant with prior literature linking food environments and food insecurity to adverse cardiovascular outcomes. Several prior studies reinforce the association observed in our analysis. A study by Sun et al. demonstrated that food insecurity is significantly associated with increased risk of both all-cause and cardiovascular mortality, with individuals experiencing marginal and very low food security showing higher CVD mortality risk.[15] Similarly, Parekh et al. reported that U.S. counties with the least healthy food environments had substantially higher cardiovascular mortality rates compared to those with the healthiest environments (704.3 vs 598.6 deaths per 100,000 population), highlighting the population-level impact of food access.[19] Consistent with these findings, Fu et al. and Gundersen et al. showed that food insecurity is associated with increased mortality and poorer diet quality, with a substantial proportion of deaths attributable to cardiovascular causes.[16, 17] Furthermore, Kelli et al. demonstrated that residing in food deserts is associated with a higher risk of adverse cardiovascular outcomes among patients with coronary artery disease, with area-level income emerging as a stronger determinant than physical access alone.[18]

However, the broader food-environment literature has not been uniformly consistent. Differences across studies may reflect variation in geographic scale, food-environment definitions, outcome definitions, and covariate adjustment.[5,6,8-14] Some studies focused on geographic food access alone have shown weaker or more context-dependent associations with cardiometabolic outcomes, suggesting that proximity to food outlets may not fully capture affordability, nutritional quality, transportation barriers, purchasing behaviour, or dietary intake.[5,8,9,18] In addition, county-level cardiovascular outcomes may be influenced by healthcare access, insurance coverage, rurality, and other geographically patterned determinants that are not fully captured by food environment measures.[26,27] These discordant or heterogeneous findings support interpreting RFEI as an outlet-based structural proxy rather than a direct measure of individual exposure.

Our findings should be interpreted as a complementary, RFEI-focused extension of prior county-level food environment research. Parekh et al. previously showed that a broader Food Environment Index, incorporating food insecurity and food desert status, was associated with cardiovascular mortality among adults aged 45 years and older in U.S. counties.[19] The present study asks a narrower question: whether retail outlet composition itself, measured using RFEI-based ratios, is associated with age-adjusted cardiac mortality. This distinction matters because food insecurity, geographic food access, and retail outlet composition represent overlapping but non-identical dimensions of the food environment. By showing that RFEI remained positively associated with cardiac mortality after adjustment for food insecurity, socioeconomic conditions, behavioral and cardiometabolic covariates, and spatial-error dependence, our analysis suggests that retail outlet composition may provide additional ecological information within the broader county food-environment framework. This does not establish causality, but it helps refine existing knowledge by identifying outlet composition as a measurable county-level structural correlate that warrants further longitudinal and quasi-experimental study.

Consistent with this literature, our national ecological analysis found that counties with less favourable retail food environments had higher cardiac mortality. The association was modest in unadjusted analyses but persisted after multivariable adjustment and remained positive after accounting for spatially structured residual dependence, although the spatial error model attenuated the magnitude of the estimate. Together with prior evidence linking food insecurity, unhealthy food environments, and food deserts to cardiovascular outcomes, these findings suggest that the retail food environment may capture an important component of the broader socioeconomic and structural context in which cardiovascular risk develops. However, because RFEI

is an area-level exposure and this study is ecological and cross-sectional, it should be interpreted as a county-level structural correlate of cardiac mortality rather than as evidence of an individual-level or causal effect.

The incremental contribution of this study lies in evaluating this relationship using a simple and reproducible outlet-ratio measure within a national county-level framework, while also examining the effect of alternative superstore classification, explicitly considering covariates that may lie partly on the causal pathway, and incorporating spatial sensitivity analyses. These features allow us to test the robustness of the observed association under several analytic specifications. Accordingly, this study extends rather than replaces prior county-level food-environment research by providing additional evidence that the composition of the local retail food environment is associated with geographic variation in cardiac mortality.

The spatial sensitivity analysis materially improved interpretation of the findings. Residual spatial dependence was present after conventional OLS modeling, indicating that geographically clustered unmeasured factors contributed to county-level variation in cardiac mortality. After fitting a spatial error model, the association between RFEI and mortality was attenuated but remained positive and statistically significant. This suggests that part of the OLS association reflected spatially structured contextual factors, while the RFEI association was not fully explained by spatial dependence. These findings remain ecological and should not be interpreted causally.

The inverse obesity coefficient in the fully adjusted model should not be interpreted as protective. In simpler models, obesity prevalence was positively associated with cardiac mortality; the coefficient reversed direction only after inclusion of multiple correlated behavioral, cardiometabolic, socioeconomic, and demographic variables. This pattern is consistent with suppression, ecological artifact, or conditional overadjustment rather than a biologically protective effect.

Importantly, while the observed associations are consistent across multiple model specifications, the cross-sectional ecological design precludes causal inference. Consistency across overlapping RFEI definitions should not be interpreted as independent replication.

Future research should prioritise longitudinal analyses, natural experiments, population-weighted approaches, and spatially explicit models that incorporate healthcare access, rurality, and other geographically patterned determinants of the relationship between food environments and cardiovascular outcomes. Specifically, future longitudinal studies could evaluate whether changes in county-level retail food outlet composition are associated with subsequent changes in cardiac mortality, compared with counties without similar changes, while accounting for socioeconomic conditions and healthcare access.

LIMITATIONS

This study’s ecological design means that findings at the county level cannot be directly applied to individuals. Associations observed in aggregated data may not reflect individual-level behaviours, and inferring such relationships risks ecological fallacy. The observed patterns suggest statistical associations at the county scale, not definitive causal pathways.

Several data limitations should be noted. Missing values and potential misclassification of food outlets are possible, particularly given our reliance on the North American Industry Classification System (NAICS) codes to categorise establishments.[25] Descriptive comparisons indicated differences between included and excluded counties in several observed characteristics, particularly RFEI and grocery access. However, substantial missingness among excluded counties limited comprehensive comparison and may have introduced selection bias.

The Retail Food Environment Index (RFEI), while widely used, is a crude proxy for the nutritional landscape. It quantifies outlet types but does not account for within-category variation in product quality, pricing, marketing, or product placement, nor does it reflect actual purchasing or consumption behaviours. Factors such as local food prices, cultural dietary norms, targeted marketing, and policy measures (e.g., menu labelling laws) can influence diet and health but are unmeasured in our dataset.

Our models did not include direct measures of healthcare access (e.g., physician density, hospital beds per capita, travel distance to facilities) due to a lack of complete national data for the study period. Healthcare availability is a known determinant of cardiac mortality, and differences across counties, such as high uninsured rates, absence of cardiologists, or lack of hospital infrastructure, could partially explain the observed variation.[26,27] We acknowledge this as an important unmeasured confounder and recommend that future work incorporate healthcare access variables or conduct targeted sensitivity analyses.

Another consideration is temporal alignment. Mortality data covered 2018-2020, while food environment and covariate measures were obtained from the USDA Food Environment Atlas 2020 data file. The Atlas release year does not necessarily indicate identical measurement years for every underlying variable.

Several covariates included in adjusted models, such as diabetes, obesity, and physical inactivity, may lie along potential causal pathways. As a result, multivariable estimates represent conditional associations rather than total effects. Sensitivity analyses excluding these variables yielded consistent findings.

Racial and ethnic composition variables are compositional and correlated with socioeconomic and geographic structure. Percentage White was omitted from the primary model, VIFs were reported, and race/ethnicity coefficients were not interpreted as individual-level or biological effects.

Although spatial dependence was addressed using a spatial error model, residual geographic confounding remains possible. The significant spatial error parameter suggests that unmeasured spatially clustered factors, such as healthcare access, environmental exposures, regional policy differences, rurality, or built-environment characteristics, may contribute to county-level mortality variation. Therefore, the spatial model should be interpreted as a sensitivity analysis improving inference for spatial dependence, not as evidence of causality. Because a formal post-spatial-error residual Moran diagnostic was unavailable, the extent of remaining residual spatial autocorrelation could not be established.

This study is conceptually related to prior county-level food environment research, particularly the work by Parekh et al.,[19] therefore, its contribution should be interpreted as incremental rather than wholly novel. The added value is the narrower RFEI-focused assessment of retail outlet composition and the finding that this measure remained associated with cardiac mortality after adjustment for food insecurity, socioeconomic conditions, cardiometabolic covariates, and spatial-error dependence.

Despite these constraints, the associations between RFEI and cardiac mortality remained positive across OLS, mediator-exclusion, log-transformed, and spatial error analyses. The findings remain hypothesis-generating.

Conclusion

Our findings underscore the critical role of place, and specifically, the local food environment, in shaping cardiovascular health outcomes across U.S. counties. Higher RFEI values were associated with higher age-adjusted cardiac mortality at the county level. Because this was a cross-sectional ecological analysis with residual spatial dependence and unmeasured confounding, the findings should be interpreted as evidence of a population-level structural correlate rather than causal evidence that modifying retail food environments would reduce mortality.

Longitudinal, quasi-experimental, population-weighted, and spatially explicit studies are needed to evaluate whether changes in retail food environments contribute to changes in cardiovascular outcomes. The present findings support continued study of retail food environments within broader structural and geographic determinants of cardiac mortality

Data availability

The study used publicly available county-level data from CDC WONDER and the USDA Food Environment Atlas. Supplementary Table 1 lists the data sources, variables, and available reference periods.

References

1. National Center for Health Statistics. FastStats: leading causes of death. Centers for Disease Control and Prevention. Available from: https://www.cdc.gov/nchs/fastats/leading-causes-of-death.htm. Cited May 21, 2025.
2. Martin SS, Aday AW, Allen NB, Almarzooq ZI, Anderson CAM, Arora P, et al. 2025 heart disease and stroke statistics: a report of US and global data from the American Heart Association. *Circulation*. 2025;151(8):e41–e660.
3. Economic Research Service. Food security in the U.S.: key statistics and graphics. U.S. Department of Agriculture. Available from: https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/key-statistics-graphics. Cited May 21, 2025.
4. Chang R, Javed Z, Taha M, Yahya T, Valero-Elizondo J, Brandt EJ, et al. Food insecurity and cardiovascular disease: current trends and future directions. Am J Prev Cardiol. 2022;9:100303.
5. Rideout K, Mah CL, Minaker L. Food environments: an introduction for public health practice. Vancouver: National Collaborating Centre for Environmental Health; 2015. Available from: https://ncceh.ca/sites/default/files/Food_Environments_Public_Health_Practice_Dec_2015.pdf. Cited May 21, 2025.
6. Downs SM, Ahmed S, Fanzo J, Herforth A. Food environment typology: advancing an expanded definition, framework, and methodological approach for improved characterization of wild, cultivated, and built food environments toward sustainable diets. Foods. 2020;9(4):532.
7. Schönfeldt HC, Pretorius B. Agriculture and food systems for improved nutrition. Adv Food Secur Sustain. 2018; 3:53-68.
8. Walker RE, Keane CR, Burke JG. Disparities and access to healthy food in the United States: a review of food deserts literature. Health Place. 2010;16(5):876-884.
9. Beaulac J, Kristjansson E, Cummins S. A systematic review of food deserts, 1966–2007. Prev Chronic Dis. 2009;6(3): A105.
10. Cooksey-Stowers K, Schwartz MB, Brownell KD. Food swamps predict obesity rates better than food deserts in the United States. Int J Environ Res Public Health. 2017;14(11):1366.
11. Antrum CJ, Waring ME, Cohen JFW, Stowers KC. Within-store fast food marketing: the association between food swamps and unhealthy advertisement. Prev Med Rep. 2023; 35:102349.
12. Pineda E, Stockton J, Mindell JS. The Retail Food Environment Index and its association with dietary patterns, body mass index, and socioeconomic position: a multilevel assessment in Mexico. PLOS Glob Public Health. 2024;4(10):e0003819.
13. County Health Rankings & Roadmaps. Food Environment Index. University of Wisconsin Population Health Institute. Available from: https://www.countyhealthrankings.org/health-data/community-conditions/health-infrastructure/health-promotion-and-harm-reduction/food-environment-index. Cited May 21, 2025.
14. Glover B, Mao L, Hu Y, Zhang J. Enhancing the Retail Food Environment Index with neighborhood commuting patterns: a hybrid human-environment measure. Int J Environ Res Public Health. 2022;19(17):10798.
15. Sun Y, Liu B, Rong S, Du Y, Xu G, Snetselaar LG, et al. Food insecurity is associated with cardiovascular and all-cause mortality among adults in the United States. J Am Heart Assoc. 2020;9(19):e014629.
16. Fu P, Wen J, Duan X, Hu X, Chen F, Yuan P. Association between adult food insecurity and mortality among adults aged 20–79 years with diabetes: a population-based retrospective cohort study. Diabet Med. 2024;41(4):e15268.
17. Gundersen C, Tarasuk V, Cheng J, de Oliveira C, Kurdyak P. Food insecurity status and mortality among adults in Ontario, Canada. PLoS One. 2018;13(8):e0202642.
18. Kelli HM, Kim JH, Tahhan AS, Liu C, Ko YA, Hammadah M, et al. Living in food deserts and adverse cardiovascular outcomes in patients with cardiovascular disease. J Am Heart Assoc. 2019;8(4):e010694.

19. Parekh T, Xue H, Al-Kindi S, Nasir K, Cheskin LJ, Cuellar AE. Food environment quality and cardiovascular disease mortality in the United States: a county-level analysis from 2017 to 2019. J Gen Intern Med. 2024;39(2):176-185.
20. Opendatasoft. Counties: United States of America. Available from: https://public.opendatasoft.com/explore/dataset/georef-united-states-of-america-county/. Cited May 21, 2025.
21. Centers for Disease Control and Prevention. Interactive Atlas of Heart Disease and Stroke: tables. Centers for Disease Control and Prevention. Available from: https://nccd.cdc.gov/dhdspatlas/reports.aspx. Cited May 21, 2025.
22. Centers for Disease Control and Prevention. Multiple Cause of Death, 2018–2023, Single Race Request Form. CDC WONDER. Available from: https://wonder.cdc.gov/controller/datarequest/D157. Cited May 21, 2025
23. Economic Research Service. Food Environment Atlas: data access and documentation downloads. U.S. Department of Agriculture. Available from: https://www.ers.usda.gov/data-products/food-environment-atlas/data-access-and-documentation-downloads. Cited May 21, 2025
24. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology statement: guidelines for reporting observational studies. Lancet. 2007;370(9596):1453-1457.
25. U.S. Census Bureau. North American Industry Classification System. U.S. Census Bureau. Available from: https://www.census.gov/naics/. Cited May 21, 2025.
26. American College of Cardiology. Almost half of US counties have no cardiologists despite higher prevalence of CV risk factors, mortality. Washington: American College of Cardiology; 2024 Jul 8. Available from: https://www.acc.org/About-ACC/Press-Releases/2024/07/08/18/25/Almost-Half-of-US-Counties-Have-No-Cardiologists-Despite-Higher-Prevalence-of-CV-Risk-Factors-Mortality. Cited May 21, 2025.
27. Ibrahim R, Habib A, Terrani K, Ravi S, Takamatsu C, Salih M, et al. County-level variation in healthcare coverage and ischemic heart disease mortality. PLoS One. 2024;19(1):e0292167.

Figures

Figure 1. Flow Diagram of County Selection and Exclusions for the Nationwide Ecological Study of Retail Food Environment and Cardiac Mortality in the United States

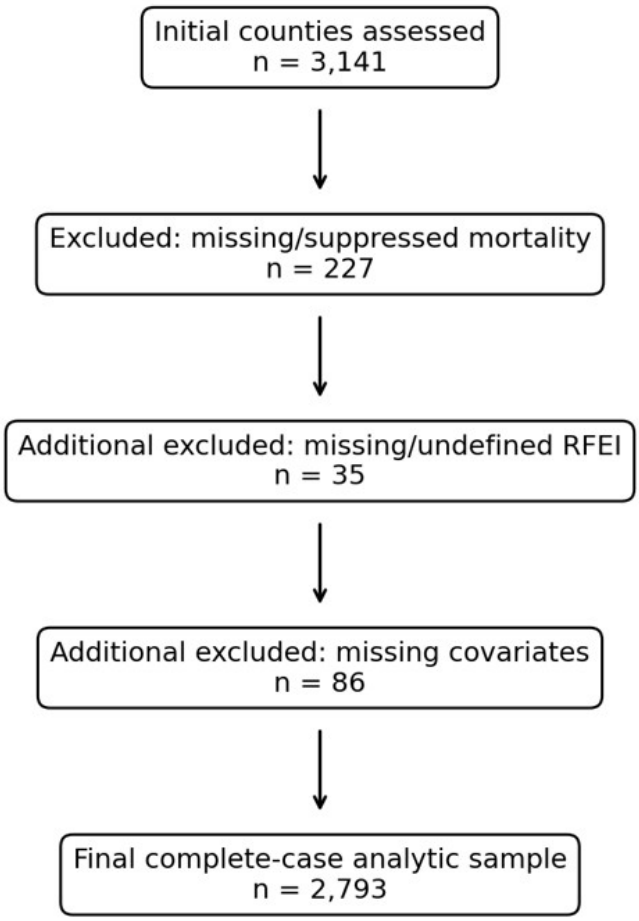
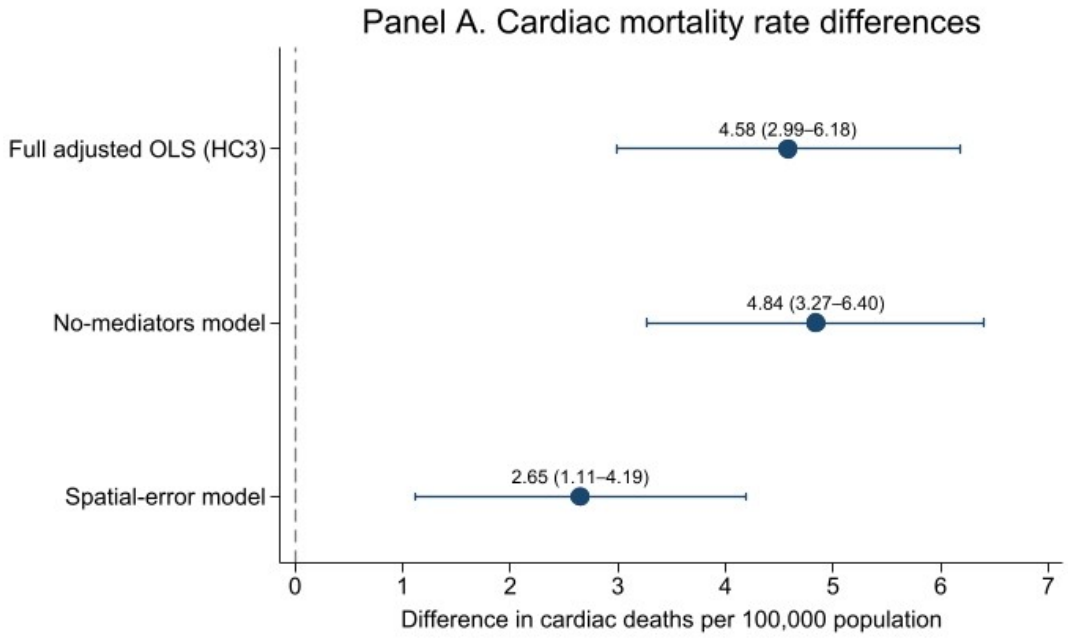


Figure 2. Sensitivity-analysis estimates for Retail Food Environment Indices and Cardiac Mortality in 2,793 U.S. Counties

(A) Adjusted differences in county-level cardiac mortality rates (deaths per 100,000 population per 1-SD higher RFEI) from the full OLS, no-mediators, and spatial-error models.



(B) Percentage differences in cardiac mortality per 1-SD higher RFEI, RFEI1, or RFEI2 from log-transformed outcome models. Points represent effect estimates, and horizontal lines indicate 95% confidence intervals.

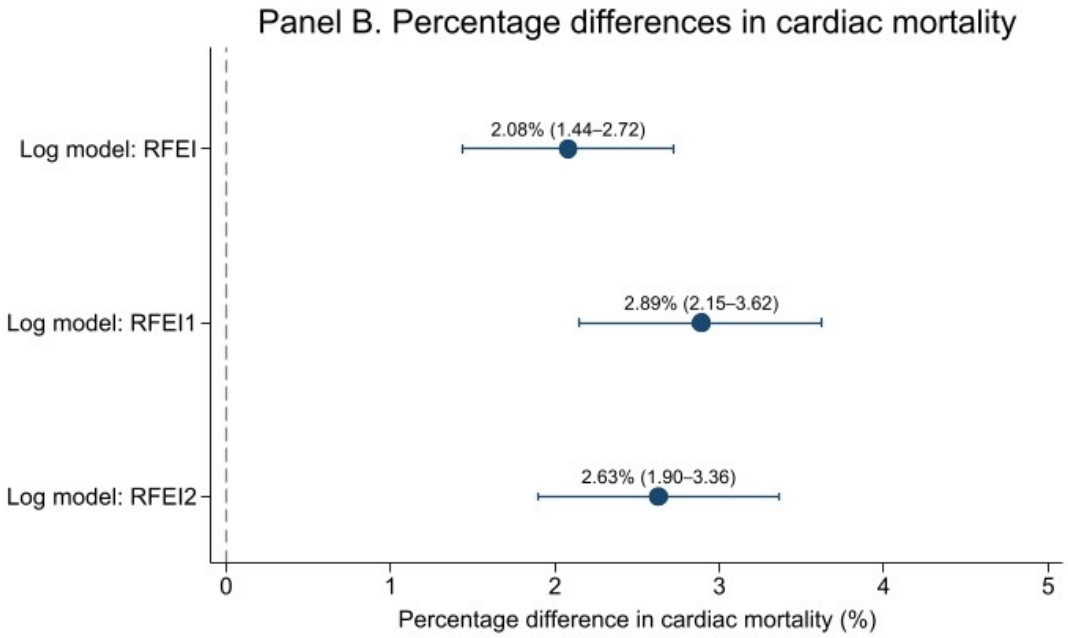
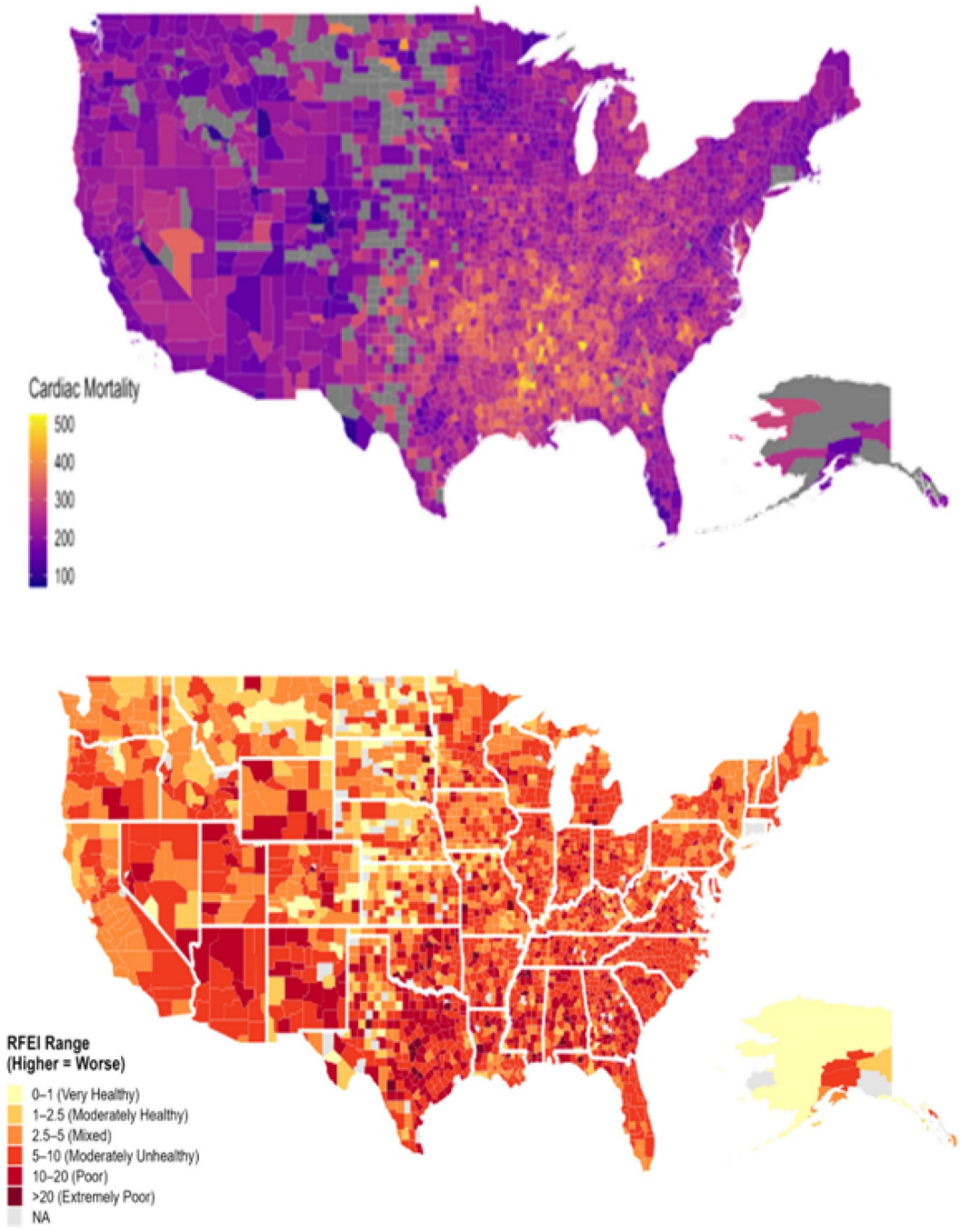


Figure 3. Geographic Distribution of Age-Adjusted Cardiac Mortality and Retail Food Environment Index Across U.S. Counties



Notes: National county-level choropleth maps illustrating geographic variation in the Retail Food Environment Index (RFEI) and age-adjusted cardiac mortality, expressed as deaths per 100,000 population. Higher RFEI values indicate a greater relative concentration of food outlets conventionally classified as less healthy. Retail food outlet measures primarily represent 2016 data, whereas cardiac mortality rates represent 2018–2020.

Table 1. County-Level Cardiac Mortality, Retail Food Environment Indices, and Population Characteristics Across 2,793 U.S. Counties

Variable	Mean	SD
RFEI	7.01	4.34
RFEI1	4.64	2.70
RFEI2	4.15	2.23
Cardiac mortality (per 100,000)	246.98	58.46
% White	78.58	19.46
% Black	9.12	14.72
% Hispanic	8.00	12.85
% Asian	1.17	2.45
% Native American	1.48	5.74
% Pacific Islander	0.07	0.39
Median household income ($)	48,633.74	12,443.52
Poverty rate (%)	16.30	6.31
Low grocery-store access (%)	20.19	14.13
SNAP participation (%)	13.00	3.15
Food insecurity (%)	12.92	2.24
Diabetes prevalence (%)	8.65	1.56
Obesity prevalence (%)	28.69	6.34
Physical inactivity (%)	19.56	4.00
Smoking prevalence (%)	18.19	3.92

Notes: Values represent county-level means and standard deviations.

RFEI = Retail Food Environment Index; SNAP = Supplemental Nutrition Assistance Program

Table 2. Correlation Between Food Environment Indices and Age-Adjusted Cardiac Mortality Across 2,793 U.S. Counties

Dataset- USDA Atlas Edition	Exposure	Pearson Coefficient (r)	p-value
2020	RFEI	0.13	<0.001
2020	RFEI1	0.25	<0.001
2020	RFEI2	0.24	<0.001

Notes: Pearson correlation coefficients (r) shown. The 2020 designation refers to the USDA Food Environment Atlas edition rather than a common measurement year for all underlying indicators. Cardiac mortality data cover 2018-2020.

Table 3. Univariable Associations Between Retail Food Environment Indices and Age-Adjusted Cardiac Mortality Across 2,793 U.S. Counties

Exposure	β (95% CI)	p-value	R^2
RFEI	1.73 (1.24-2.23)	<0.001	0.017
RFEI1	5.33 (4.55-6.10)	<0.001	0.061
RFEI2	6.38 (5.44-7.33)	<0.001	0.059

Notes: β represents the change in cardiac mortality (per 100,000 population) per 1-unit increase in exposure.

Table 4. Multivariable Associations Between Retail Food Environment, Socioeconomic, Demographic, and Behavioral Factors and Cardiac Mortality Across 2,793 U.S. Counties

Variable	β (95% CI), HC3	p-value
RFEI	4.58 (2.99–6.18)	<0.001
Median income	−7.11 (−10.41 to −3.80)	<0.001
Poverty rate	4.65 (0.25–9.05)	0.038
SNAP participation	2.96 (0.95–4.96)	0.004
Food insecurity	4.22 (2.49–5.96)	<0.001
Grocery access	−0.28 (−2.24 to 1.69)	0.783
% Black	8.04 (5.31–10.78)	<0.001
% Hispanic	−2.07 (−4.21 to 0.07)	0.058
% Asian	−1.47 (−3.37 to 0.43)	0.128
% Native American	−2.93 (−5.23 to −0.63)	0.013
% Pacific Islander	1.07 (−0.52 to 2.65)	0.187
Diabetes prevalence	3.68 (0.90–6.47)	0.010
Obesity prevalence	−6.02 (−8.50 to −3.54)	<0.001
Smoking prevalence	20.98 (17.35–24.60)	<0.001
Physical inactivity	6.18 (3.45–8.92)	<0.001

Notes: All continuous predictors were z-standardized. Beta values represent differences in cardiac mortality per 100,000 population per 1-SD higher predictor. Percentage White was omitted due to compositional collinearity. Race/ethnicity coefficients are contextual county-level controls and are not interpreted as individual-level or biological effects. HC3 Heteroskedasticity-consistent standard errors were used (n=2793).

Table 5. Sensitivity Analysis of Retail Food Environment Indices and Cardiac Mortality Across 2,793 U.S. Counties

Analysis	Estimate	95% CI	p value
Full adjusted OLS (HC3)	4.58	2.99-6.18	<0.001
No-mediators model	4.84	3.27-6.40	<0.001
Spatial error model	2.65	1.11-4.19	0.001
Log model: RFEI	2.08%	1.44%-2.72%	<0.001
Log model: RFEI1	2.89%	2.15%-3.62%	<0.001
Log model: RFEI2	2.63%	1.90%-3.36%	<0.001

Notes: OLS and spatial-error estimates represent differences in cardiac deaths per 100,000 population per 1-SD higher predictor. The full adjusted OLS model used HC3 standard errors. Log-transformed outcome models used Stata's vce (robust) estimator, with percentage differences calculated as 100 × [exp(β) - 1] and delta-method confidence intervals. The spatial-error model was estimated using generalized spatial two-stage least squares. The no-mediators model included 2,793 counties and used Stata's vce(robust) heteroskedasticity-consistent variance estimator.

Table 6. Model Diagnostics and Spatial Dependence in Adjusted County-Level Cardiac Mortality Models

Diagnostic	Result
Main analytic sample	2,793 counties
Spatial analytic sample	2,793 counties
Spatial weights matrix dimension	2,793 × 2,793
Full adjusted OLS R^2	0.502627
Full adjusted OLS adjusted R^2	0.499941
Maximum VIF	5.44
Mean VIF	2.31
RFEI VIF	1.09
Breusch-Pagan/Cook-Weisberg chi-square	186.97
Breusch-Pagan/Cook-Weisberg p value	<0.001
Moran test chi-square	368.27
Moran test p value	<0.001
Spatial error parameter rho	0.47 (95% CI, 0.43-0.51)
Spatial error parameter p value	<0.001
Counties with no contiguous neighbors	12
Mean neighbors per county	5.51

Notes: VIF = variance inflation factor. Spatial analyses used a row-normalized county-contiguity weights matrix among counties linked to boundary data. Counties without contiguous neighbors were retained in the spatial analytic sample and assigned zero neighbor weights. The reported Moran test was conducted following conventional OLS. A formal post-spatial-error residual Moran diagnostic was not available through the generalized spatial two-stage least-squares workflow; therefore, residual spatial autocorrelation after spatial-error estimation was not formally assessed.

Supplementary Material

Supplementary Table 1. Source variables used to construct RFEI indices and analytic covariates

Domain	Variable description	Source variable code/ analytic variable	Data source	Role in analysis
Food environment	Grocery stores per 1,000 population	GROCPTH16	USDA Food Environment Atlas 2020 edition	RFEI denominator; RFEI1 denominator; RFEI2 denominator
Food environment	Supercenters and warehouse club stores per 1,000 population	SUPERCPTH16	USDA Food Environment Atlas 2020 edition	RFEI1 numerator; RFEI2 denominator
Food environment	Convenience stores per 1,000 population	CONVSPTH16	USDA Food Environment Atlas 2020 edition	RFEI numerator; RFEI1 numerator; RFEI2 numerator
Food environment	Specialized food stores per 1,000 population	SPECSPTH16	USDA Food Environment Atlas 2020 edition	RFEI1 denominator; RFEI2 denominator
Food environment	Fast-food restaurants per 1,000 population	FFRPTH16	USDA Food Environment Atlas 2020 edition	RFEI numerator; RFEI1 numerator; RFEI2 numerator
Food environment	Farmers markets per 1,000 population	FMRKTPTH18	USDA Food Environment Atlas 2020 edition	RFEI1 denominator; RFEI2 denominator
Calculated index	Retail Food Environment Index	RFEI	Calculated from USDA component variables	Primary exposure
Calculated index	Expanded Retail Food Environment Index 1	RFEI1	Calculated from USDA component variables	Sensitivity exposure; supercenters classified as less favorable outlets
Calculated index	Expanded Retail Food Environment Index 2	RFEI2	Calculated from USDA component variables	Sensitivity exposure; supercenters classified as healthier food-access outlets
Demographic covariate	Percent White population	pct_white	USDA Food Environment Atlas 2020 edition	Covariate
Demographic covariate	Percent Black population	pct_black	USDA Food Environment Atlas 2020 edition	Covariate
Demographic covariate	Percent Hispanic population	pct_hispanic	USDA Food Environment Atlas 2020 edition	Covariate
Demographic covariate	Percent Asian population	pct_asian	USDA Food Environment Atlas 2020 edition	Covariate
Demographic covariate	Percent American Indian or Alaska Native population	pct_native	USDA Food Environment Atlas 2020 edition	Covariate
Demographic covariate	Percent Hawaiian or Pacific Islander population	pct_pacific	USDA Food Environment Atlas 2020 edition	Covariate
Socioeconomic covariate	Median household income	median_income	USDA Food Environment Atlas 2020 edition	Covariate
Socioeconomic covariate	Poverty rate	poverty_rate	USDA Food Environment Atlas 2020 edition	Covariate

Domain	Variable description	Source variable code/ analytic variable	Data source	Role in analysis
Food-access covariate	Population with low access to stores	grocery_access	USDA Food Environment Atlas 2020 edition	Covariate
Socioeconomic covariate	SNAP participants, percent of population	snap_pct	USDA Food Environment Atlas 2020 edition	Covariate
Food-security covariate	Household food insecurity	food_insecurity	USDA Food Environment Atlas 2020 edition	Covariate
Outcome	Age-adjusted cardiac mortality rate per 100,000 population	cardiac_mortality	CDC data, 2018-2020	Primary outcome
Cardiometabolic covariate	Age-adjusted diabetes estimate among adults aged 20 years or older	dm_pct	CDC data (reference period unspecified)	Covariate / potential mediator
Cardiometabolic covariate	Age-adjusted obesity estimate among adults aged 20 years or older	obesity_pct	CDC data (reference period unspecified)	Covariate / potential mediator
Behavioral covariate	Age-adjusted physical inactivity estimate among adults aged 20 years or older	phys_inactive_pct	CDC data (reference period unspecified)	Covariate / potential mediator
Behavioral covariate	Age-adjusted smoking estimate among adults aged 18 years or older	smoking_pct	CDC data (reference period unspecified)	Covariate

RFEI was calculated as (FFRPTH16 + CONVSPTH16) / GROCPTH16.

RFEI1 was calculated as (FFRPTH16 + CONVSPTH16 + SUPERCPTH16) / (GROCPTH16 + FMRKTPTH18 + SPECSPTH16).

RFEI2 was calculated as (FFRPTH16 + CONVSPTH16) / (GROCPTH16 + FMRKTPTH18 + SPECSPTH16 + SUPERCPTH16).

Higher values indicate a greater relative concentration of outlets conventionally classified as less favorable compared with healthier food-access outlets.

Source and measurement-year note: Food-environment, demographic, socioeconomic, and food-access variables were obtained from the FOOD DATA worksheet, while cardiac mortality and behavioral/cardiometabolic variables were obtained from the CARDIAC DATA worksheet of the source workbook. The VARIABLE LIST worksheet documents the source-variable descriptions and available reference periods.

USDA retail-outlet measures correspond to 2016 (GROCPTH16, SUPERCPTH16, CONVSPTH16, SPECSPTH16, and FFRPTH16) and 2018 (FMRKTPTH18). The USDA Food Environment Atlas 2020 designation refers to the dataset edition rather than a common measurement year.

The workbook identifies demographic percentages as 2010 measures; median household income, poverty rate, and low grocery-store access as 2015 measures; SNAP participation as a 2017 measure; and household food insecurity as a 2015–2017 three-year average. Cardiac mortality covers 2018–2020. Diabetes, obesity, physical inactivity, and smoking estimates are included in the CARDIAC DATA worksheet, but their individual reference periods are not specified in the workbook's variable list.

Supplementary Table 2. Comparison of included and excluded counties using observed data

Variable	Included n	Included mean ± SD	Excluded available n	Excluded mean ± SD
RFEI	2793	7.01 ± 4.34	266	4.03 ± 3.30
RFEI1	2793	4.64 ± 2.70	284	3.23 ± 3.13
RFEI2	2793	4.15 ± 2.23	285	3.03 ± 2.80
Cardiac mortality (per 100,000)	2793	246.98 ± 58.46	121	236.78 ± 54.13
White (%)	2793	78.58 ± 19.46	348	76.22 ± 22.68
Black (%)	2793	9.12 ± 14.72	348	5.74 ± 11.40
Hispanic (%)	2793	8.00 ± 12.85	348	10.56 ± 15.51
Asian (%)	2793	1.17 ± 2.45	348	0.86 ± 2.65
Native American (%)	2793	1.48 ± 5.74	348	4.74 ± 14.99
Pacific Islander (%)	2793	0.07 ± 0.39	348	0.20 ± 2.62
Median household income ($)	2793	48,633.74 ± 12,443.52	346	48,429.34 ± 11,602.10
Poverty rate (%)	2793	16.30 ± 6.31	346	15.97 ± 7.41
Low grocery-store access (%)	2793	20.19 ± 14.13	329	47.61 ± 35.76
SNAP participation (%)	2793	13.00 ± 3.15	348	12.00 ± 3.20
Food insecurity (%)	2793	12.92 ± 2.24	348	12.19 ± 1.61
Diabetes prevalence (%)	2793	8.65 ± 1.56	280	7.83 ± 1.30
Obesity prevalence (%)	2793	28.69 ± 6.34	280	22.82 ± 5.91
Physical inactivity (%)	2793	19.56 ± 4.00	280	17.08 ± 3.47
Smoking prevalence (%)	2793	18.19 ± 3.92	280	17.63 ± 4.26

Note: The initial dataset contained 3,141 counties; 2,793 were included and 348 were excluded by complete-case criteria. The exclusion sequence was 227 counties with missing or suppressed cardiac mortality, 35 additional counties with missing or undefined RFEI indices, and 86 additional counties with missing model covariates. For each variable, descriptive statistics among excluded counties use only observed values, so denominators vary. “Excluded” means not included in the final complete-case analytic sample, not that every variable was missing. These are descriptive comparisons only; no statistical significance testing or causal interpretation is implied.

Supplementary Table 3. Pearson correlation matrix of predictors included in the fully adjusted model (n = 2,793 counties)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. RFEI	1.0000	—	—	—	—	—	—	—	—	—	—	—	—	—	—
2. Median household income	0.0042	1.0000	—	—	—	—	—	—	—	—	—	—	—	—	—
3. Poverty rate	0.0587	-0.7910	1.0000	—	—	—	—	—	—	—	—	—	—	—	—
4. SNAP participation	0.1921	-0.3437	0.4449	1.0000	—	—	—	—	—	—	—	—	—	—	—
5. Food insecurity	0.0747	-0.1885	0.1770	0.3202	1.0000	—	—	—	—	—	—	—	—	—	—
6. Low grocery-store access	0.0315	0.0907	-0.0678	-0.0808	-0.0490	1.0000	—	—	—	—	—	—	—	—	—
7. Black (%)	0.1486	-0.2817	0.5266	0.3232	0.0726	-0.0158	1.0000	—	—	—	—	—	—	—	—
8. Hispanic (%)	0.0792	0.0478	0.0868	0.0361	-0.0618	0.1143	-0.1071	1.0000	—	—	—	—	—	—	—
9. Asian (%)	-0.0198	0.4514	-0.1693	-0.1007	-0.1386	0.0110	0.0120	0.1663	1.0000	—	—	—	—	—	—
10. Native American (%)	-0.0493	-0.0787	0.1530	-0.0054	-0.0246	0.1710	-0.0943	-0.0225	-0.0406	1.0000	—	—	—	—	—
11. Pacific Islander (%)	-0.0051	0.0984	-0.0444	-0.0379	-0.1076	0.0491	-0.0378	0.0437	0.5140	0.0024	1.0000	—	—	—	—
12. Diabetes prevalence	0.1559	-0.2889	0.3948	0.2814	0.1042	0.0043	0.4507	-0.0819	-0.0311	0.1652	-0.0177	1.0000	—	—	—
13. Obesity prevalence	0.0861	0.0155	-0.0279	-0.0345	0.0840	0.0378	0.0407	-0.1638	-0.0201	0.0871	-0.0183	0.5929	1.0000	—	—
14. Smoking prevalence	-0.0019	-0.6742	0.6260	0.3770	0.2894	-0.1461	0.2312	-0.2567	-0.3909	0.2058	-0.1181	0.3800	0.1068	1.0000	—
15. Physical inactivity	0.0961	-0.2859	0.2817	0.1815	0.1862	-0.0111	0.2545	-0.1592	-0.1002	0.0930	-0.0412	0.6920	0.6873	0.4163	1.0000

Note: Values are Pearson correlation coefficients (r) calculated among 2,793 complete-case counties. The lower triangle and diagonal are shown; blank upper-triangle cells (—) duplicate values below the diagonal. RFEI = Retail Food Environment Index; SNAP = Supplemental Nutrition Assistance Program.