Uncovering Key Predictors of County-Level Mortality in the U.S.

490455715

Introduction

Aim of the research

The purpose of this study is to identify and explain the predictors of health outcomes at the county level in the United States. Specifically, it aims to analyse the key factors affecting age-adjusted mortality rates to determine which characteristics significantly influence health outcomes.

Hypothesis Explanation

This research aims to verify the hypothesis that social determinants of health have a more substantial impact on health outcomes than environmental factors. This hypothesis is based on previous studies suggesting that not only the physical environment but also social determinants, which incorporate lifestyle and social status play critical roles in determining health (World Health Organization, 2024). Additionally, the study seeks to identify which specific factors within these determinants most significantly affect mortality rates.

Importance of the Study

Understanding the factors that influence health outcomes is crucial for developing public health policies and strategies. Reducing mortality is vital for improving the overall health level of communities and enhancing the quality of life (STIEFEL et al., 2010). By evaluating the relative importance of various factors affecting health outcomes, this study provides the necessary evidence to develop effective intervention strategies.

Research Process

The research process follows these steps:

1. **Initial Analysis with Linear Regression**

   Linear regression was used to analyse the five independent variables to assess which factors have the most significant impact on premature mortality. Then, the hypothesis is tested and leads to deeper analysis.

2. **Detailed Analysis with Lasso Regression**

   Lasso regression was applied to the major independent variables identified as significant to analyse the detailed factors within each variable. This analysis helps determine which specific factors are key predictors of better health outcomes.

3. **Geographic Visualization**

   Based on the modelling results, geographic visualisation was made to identify the characteristics that explain why certain counties have higher or lower age-adjusted mortality rates. By comprehensively analysing the main factors influencing premature mortality, it aimed to propose more effective public health policies.

Literature and theory

Understanding the determinants of health outcomes is crucial for developing effective public health policies. This study hypotheses that social determinants of health have a greater impact on health outcomes than environmental factors. The following literature review explores existing research on these causal mechanisms and provides detailed evidence supporting the hypothesis.

Social Determinants of Health Factors

1. **Health Behaviours**
   Health behaviours such as smoking, physical inactivity, excessive drinking, and poor diet are well-documented determinants of health outcomes. These behaviours directly contribute to the development of chronic diseases such as heart disease and cancer (Mokdad et al., 2004). For instance, smoking is a major cause of lung cancer, significantly increasing the risk of premature death (Jha et al., 2013). Similarly, physical inactivity and poor diet contribute to obesity, which is a significant risk factor for various chronic conditions (Ng et al., 2014).

2. **Clinical Care**
   Access to and quality of clinical care are critical components of health outcomes. Preventive services, timely medical interventions, and effective management of chronic conditions can significantly reduce mortality rates (McGinnis et al., 2002). Research indicates that areas with better healthcare access and higher quality of care experience lower mortality rates (Shi & Starfield, 2001).

3. **Social and Economic Environment**
   The social and economic environment, including income, education, employment, and social support networks, plays a critical role in shaping health outcomes. Higher income and education levels are associated with better health literacy, access to healthcare, and healthier lifestyles (Adler & Ostrove, 1999). Employment provides financial stability and access to employer-based health benefits, reducing stress and promoting health (Ross & Mirowsky, 1995). These factors collectively contribute to shaping health outcomes and mitigate the effects of environmental and behavioural factors.

Environmental Condition

1. **Physical Environment**
   The environmental conditions encompass various factors such as air and water quality, housing conditions, and access to green spaces. For instance, exposure to air pollution has been linked to respiratory and cardiovascular diseases, leading to higher mortality rates (Brunekreef & Holgate, 2002). However, while these factors are critical, they may not act in isolation and often intersect with socioeconomic and behavioural determinants.

Relative Impact on Health Outcome

1. **Immediate vs. Long-Term Impact**

   - **Social Determinants of Health**
     These factors often result in immediate health outcomes. For instance, unhealthy behaviours like smoking have immediate harmful effects on the body, increasing the risk of diseases that contribute to premature mortality (Mokdad et al., 2004).

   - **Environmental Factors**
     While it is an important factor regarding overall health outcome, the impact of environmental factors tends to be more long-term. Chronic exposure to pollutants may not show immediate health effects but can lead to significant health issues over time, such as chronic respiratory conditions (Pope et al., 2002).

2. **Empirical Evidence**
    - **Lantz et al. (1998)** found that socioeconomic and behavioural factors are more significant predictors of mortality within short to medium time frames. Their study indicates that improving socioeconomic conditions and promoting healthier behaviours can lead to rapid improvements in health outcomes.
    - **Wilkinson and Marmot (2003)** argued that addressing socioeconomic inequalities and promoting healthier behaviours can result in substantial public health improvements. Their research emphasises that while environmental improvements are essential, the immediate benefits of socioeconomic and behavioural interventions are more pronounced.

Data and methodology

Data

Dataset Used

This study used the '2018 County Health Rankings Data', which includes various health indicators for each county in the United States. The data used for analysis comprises the following sheets:

- **Additional Measure Data:** Includes the dependent variable, Premature age-adjusted mortality.
- **Ranked Measure Data:** Includes the independent variables and their detailed elements.
- **Outcomes & Factors SubRankings:** Contains the county-level ranks and quartiles for the independent variables.

Key Data Transformations and Preprocessing

- Selected necessary columns from each sheet and cleaned the variable names.
- Removed rows with missing values to create a complete dataset.
- Converted independent and dependent variables to numeric types for analysis.
- Selected the independent and dependent variables for regression analysis and prepared them for modelling.

Dependent and Independent Variables

- **Dependent Variable:** Premature age-adjusted mortality.
- **Independent Variables:** Categorised into Environmental Conditions and Social Determinants of Health based on the research hypothesis.
    - **Environmental Conditions:** Physical Environment.
    - **Social Determinants of Health:** Health Behaviours, Clinical Care, Social and Economic Environment.

Explanation of Independent Variables

- **Health Behaviours:** Evaluates the impact of behaviours such as smoking, obesity, alcohol, and drug use on age-adjusted mortality rates. It includes unhealthy lifestyle habits and risky behaviours.
- **Clinical Care:** Assesses the accessibility and quality of medical services available to people. It includes factors like the distribution of healthcare providers, access to medical services, and the effectiveness of preventive care.
- **Social and Economic Environment:** Examines the impact of living conditions on health, including economic stability, unemployment rates, social safety nets, and community safety.
- **Physical Environment:** Evaluates the impact of physical living conditions on age-adjusted mortality rates. It includes factors like air and water quality, housing conditions, and transportation accessibility.

Additional Data Used

To compare regions with high and low age-adjusted mortality rates, geographic visualisation was performed. Latitude and longitude data required for this visualisation were obtained from the US shapefiles used in Group Project 2.

Methods

Models Used

The goal of this project is to evaluate the specific impact of various factors on age-adjusted mortality rate and identify key determinants. Linear and lasso regression were used to analyse continuous dependent variables and to select important variables for model interpretability.

Linear Regression

- **Principle:** Linear regression is a statistical method that uses one or more independent variables (predictors) to predict a dependent variable (response).
- **Reasonability:** The linear regression results quantitatively demonstrated the impact of independent variables on the dependent variable, allowing us to identify which factors significantly influence age-adjusted mortality (Onwuegbuzie & Daniel, 2019). Additionally, by effectively handling numerous continuous variables, the results provided clearer interpretations and insights.

Lasso Regression

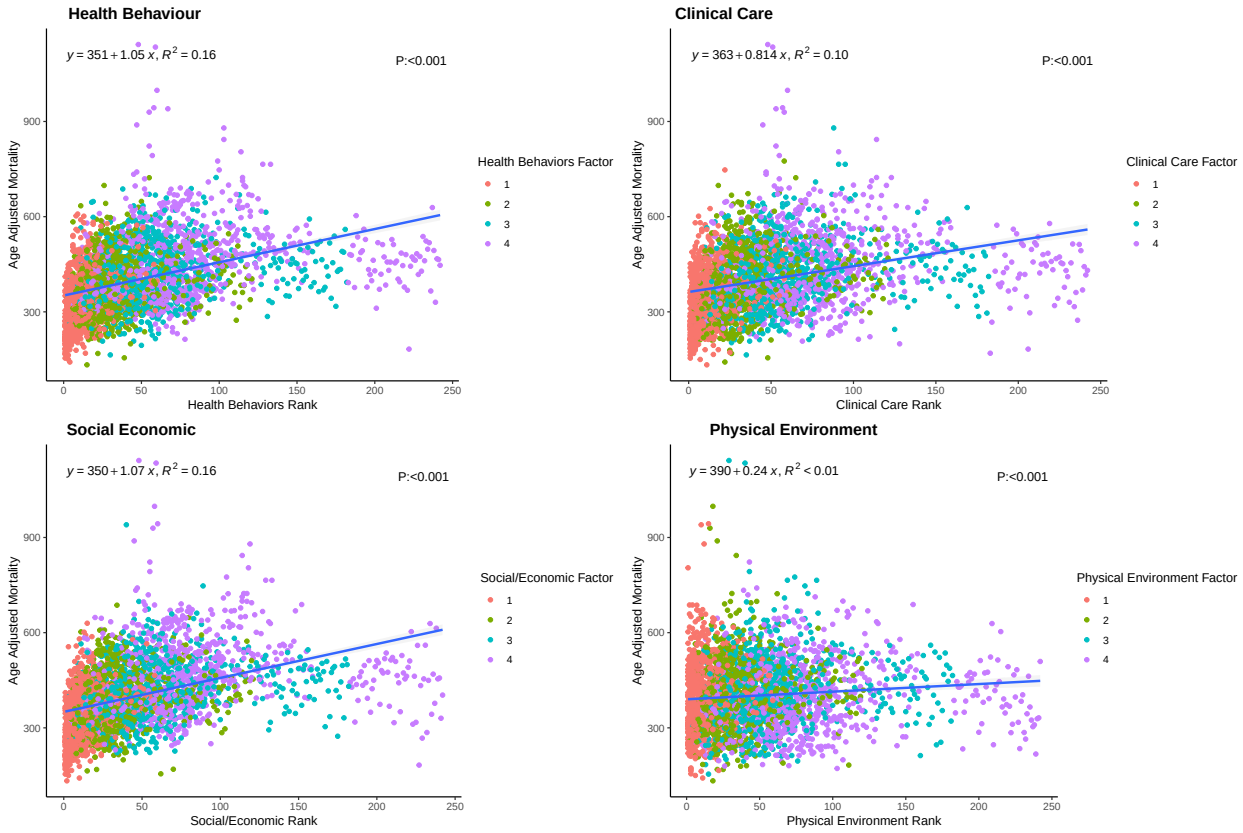
- **Principle:** Lasso regression adds an L1 regularisation penalty to the linear regression model to shrink the coefficients of less important variables to zero, effectively performing variable selection (Finch & Finch, 2019).
- **Reasonability:** Lasso regression was used to evaluate the importance of detailed factors within each independent variable and eliminate less significant variables. This helped identify the most critical predictors of health outcomes from a large number of variables and aided in model interpretation and insight generation (Finch & Finch, 2019).

Results

Project Flow / Direction

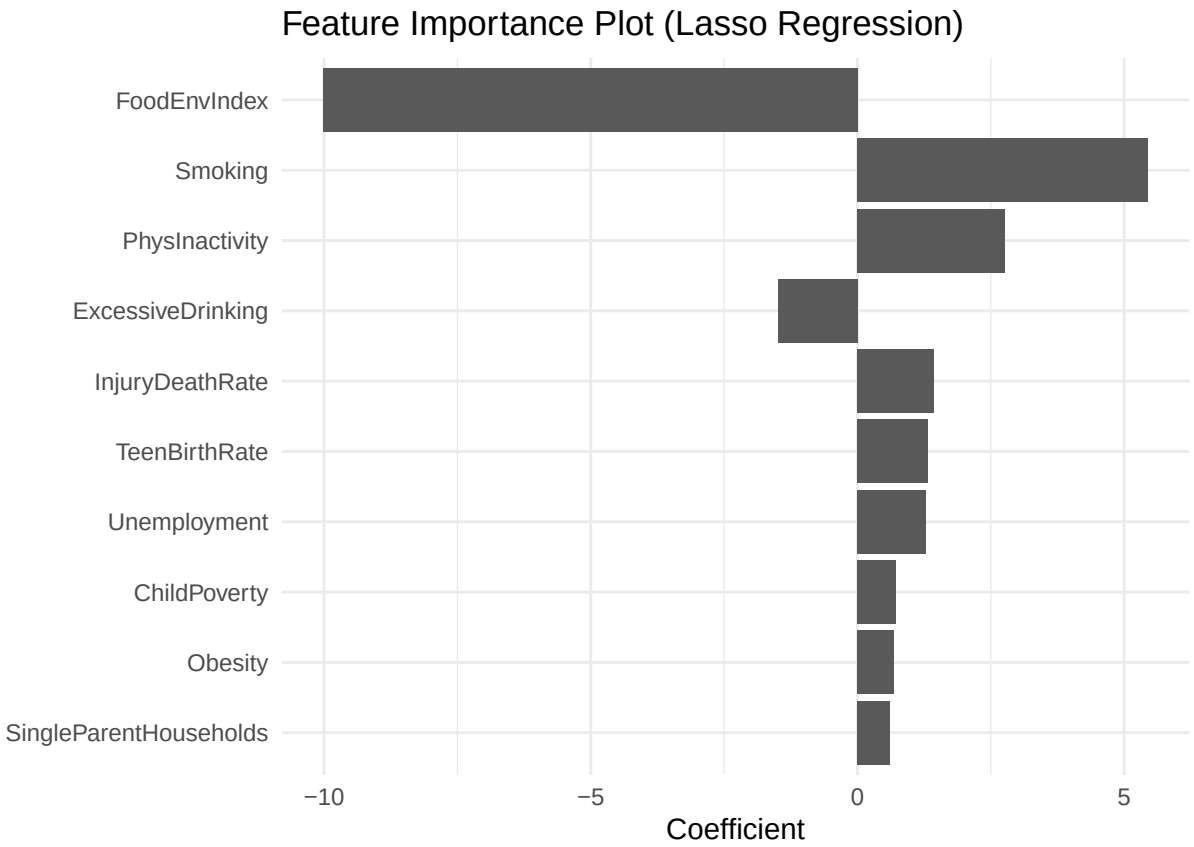
1. **Initial Analysis of Dependent and Independent Variables**
   The relationship between the dependent variable and independent variables was analysed using linear regression and visualisation. This helped assess the impact of each independent variable on the dependent variable and evaluate their importance. The initial analysis results indicate the importance of the independent variables on Age-Adjusted Mortality in the following order:

   1. Health Behaviours ($R^2 = 0.16$)
   2. Social/Economic Factors ($R^2 = 0.16$)
   3. Clinical Care ($R^2 = 0.10$)
   4. Physical Environment ($R^2 < 0.01$)



2. **Hypothesis Testing**
   Based on the initial analysis results, the hypothesis that Social Determinants of Health (Health Behaviours, Clinical Care, Social and Economic Environment) have a greater impact on Age-Adjusted Mortality Rate than Environmental Factors (Physical Environment) was validated. To analyse deeper into this hypothesis, the analysis was extended to identify and evaluate the detailed factors within the three independent variables classified under social determinants of health. This deeper analysis aimed to identify key factors for effective public health policy development and strategies to improve health outcomes.

3. **Detailed Factor Analysis Results and Insights**
   Lasso regression was used to evaluate the impact of each detailed factor on health outcomes, and visualisation helped to clearly identify the significant factors.



Key Factors

The Lasso regression analysis identified the top 10 detailed factors, with six belonging to health behaviours and four to social/economic environments. This indicates that the key predictors of county-level mortality rates are health behaviours and social/economic environments. Notably, clinical care, such as healthcare accessibility, had relatively less impact on mortality rates.

Most Influential Factor: Health Behaviours

The top four factors with the greatest impact on mortality rates all belong to health behaviours, highlighting that individuals' lifestyle choices have a direct influence on mortality rates. These factors include:

1. Food Environment Index
2. Adult Smoking (% Smoking Rate)
3. Physical Inactivity (% Physical Inactivity Rate)
4. Excessive Drinking (% Excessive Drinking Rate)

These factors demonstrate that healthy lifestyle choices significantly affect premature age-adjusted mortality rates. For instance, excessive smoking or drinking have immediate harmful effects on the body, leading to serious health issues. Additionally, maintaining a healthy diet and regular physical activity contributes to overall health.

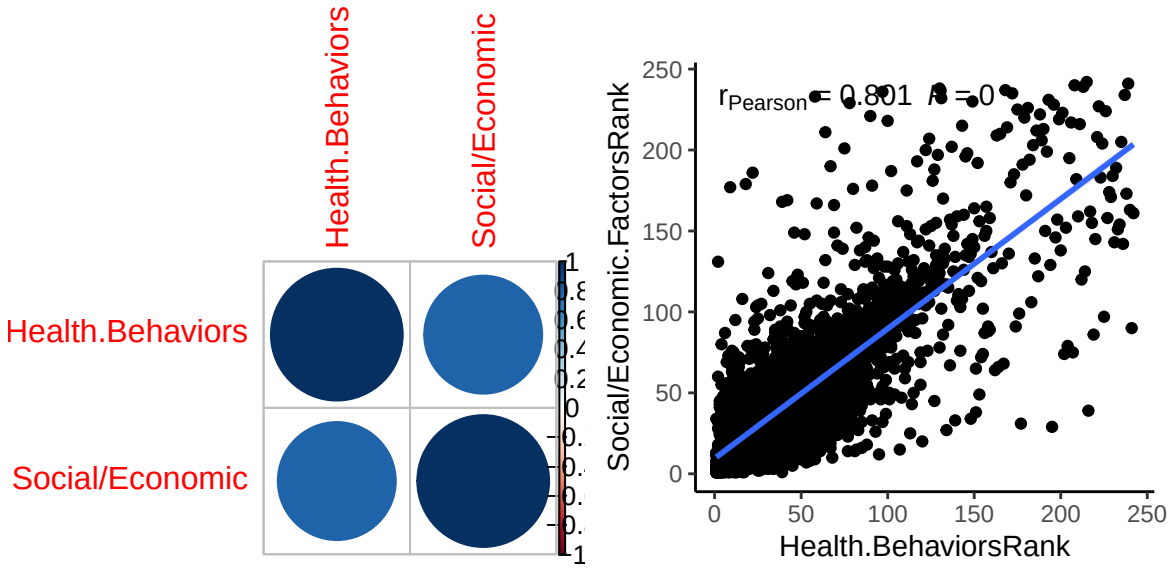
Reasons for the Relatively Lower Impact of Clinical Care

- **Degree of Impact on Mortality Rates**
  Health behaviours and social/economic factors have immediate and direct effects on mortality rates. In contrast, clinical care focuses primarily on treating existing health issues, which may result in a relatively lower impact on mortality rates.
- **Limitations of Preventive Care**
  Preventive care plays a crucial role in health management but has limitations in completely resolving existing health issues. For example, factors like insurance coverage and the number of dentists, while important for health management, may not have a direct and substantial impact on mortality rates. Therefore, these factors have a relatively minor effect on premature age-adjusted mortality rates.

Relationship Between Health Behaviours and Socioeconomic Factors

While health behaviours and social/economic environments independently influence premature mortality, their interaction can amplify their impact (Lantz et al., 1998). For instance, the ability to maintain healthy lifestyle habits is often determined by socioeconomic factors. Therefore, analysing the correlation between these two variables can provide insights and design more effective intervention strategies.

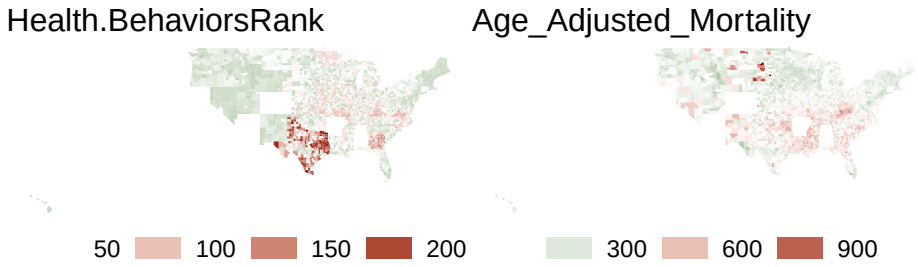
The analysis results indicated a strong correlation between them, with a Pearson's correlation coefficient of 0.801, demonstrating a significant relationship (p-value = 0).



4. **Geographic Visualization**
   Based on the analysis that health behaviours have the greatest impact on age-adjusted mortality rate among all independent variables, the spatial distribution of health behaviours rank and age-adjusted mortality across US counties was visually compared and analysed. The results showed that regions with poor performance on both maps were concentrated in the southern and southeastern United States. This indicates a correlation between poor health behaviours and high age-adjusted mortality in these regions.

   A notable example of poor performance in both health behaviours rank and age-adjusted mortality is Mississippi (e.g., Bolivar, Holmes, Leflore counties). The high mortality rates and poor health behaviours in the southern and southeastern US are due to various factors. First, the southern region has lower income levels and higher poverty rates compared to other regions. Economic inequality directly affects health, as lower income limits access to medical services or healthy food. Additionally, the southern region has many 'food desert' areas where it is difficult for residents to access healthy food, leading to increased consumption of unhealthy food and higher risks of obesity and related diseases (Walker et al., 2010).

   Addressing these issues requires targeted interventions and efforts at the county level, particularly focusing on reducing income inequality, expanding educational opportunities, improving access to medical services, and promoting healthy behaviours based on the analysis of why certain regions have poor health outcomes.



5. **Comprehensive Analysis and Insights**
   The comprehensive analysis strongly supports the hypothesis that social determinants of health have a greater impact on health outcomes than environmental factors. Among these determinants, health behaviours emerged as the most significant factor affecting premature mortality rates. Specifically, healthy eating habits, smoking, excessive drinking, and other lifestyle choices directly and significantly influence health outcomes. This finding suggests that public health policies should prioritise improving health behaviours.

   Additionally, while it is crucial to develop independent improvement strategies for health behaviours, it is also essential to adopt a comprehensive approach that considers the interaction between health behaviours and the next most influential factor, the social and economic environment.

   By conducting this analysis, specific and practical strategies can be developed to reduce mortality rates, thereby enhancing the overall health of the population and reducing early mortality rates.

   Recommendations

   1. **Health Behavior Improvement Programs**
      - **Smoking Cessation Campaigns:** Strengthen smoking cessation campaigns and support programs in areas with high smoking rates.
      - **Promotion of Physical Activity:** Expand exercise facilities and offer accessible exercise programs to residents.
      - **Alcohol Moderation Education:** Enhance education and counselling services on alcohol moderation in areas with high excessive drinking rates.
   2. **Improvement of Socioeconomic Environment:**
      - **Economic Support:** Provide economic support and job creation programs in areas with high unemployment rates and low income levels.
      - **Expansion of Educational Opportunities:** Expand educational opportunities and strengthen vocational training programs in areas with low education levels.
   3. **Establishment of Integrated Public Health Policies:**
      - **Policies Considering Both Health Behaviours and Socioeconomic Environment:** Develop public health policies that integrate improvements in health behaviours and socioeconomic environments (Lantz et al., 1998). For example, introduce programs to improve access to healthy food in low-income areas and enhance health education to promote healthy behaviours.
   4. **Region-Specific Approaches:**
      - **Tailored Policies for Different Regions:** Develop customised policies that reflect the characteristics of each region by analysing data to identify specific causes of mortality in different counties.

Conclusion

This study indicated that social determinants of health have a more significant impact on health outcomes than environmental factors. Specifically, health behaviours and social/economic environment emerged as the most critical determinants of mortality rates. The result aligns with existing literature emphasising the importance of social determinants of health in determining health outcomes. Studies by Mokdad et al. (2004) and Lantz et al. (1998) support findings that improving socioeconomic conditions and promoting healthier behaviours can lead to rapid improvements in health outcomes. To effectively reduce age-adjusted mortality rates, public health policies should prioritise interventions targeting health behaviours and socioeconomic factors. Furthermore, a comprehensive approach that integrates environmental improvements with these social determinants is recommended to achieve sustainable health benefits.

Limitations

- **Scope of Variables:** The analysis is limited to the variables present in the available dataset. Other significant factors influencing mortality rates and health outcomes, such as genetic predispositions, specific local health policies, and community health programs, were not considered. This limitation means that the study's conclusions are based only on the included variables and may not fully capture all determinants of health outcomes.
- **Temporal Inconsistency:** In the provided dataset, the years of measurement differ for each element. Without clear justification for these varying measurement periods, it can be challenging to ensure the accuracy and comparability of the impact and outcomes of different factors. This temporal inconsistency may introduce bias and affect the reliability of the study's findings.

Bibliography

ADLER, N. E., & OSTROVE, J. M. (1999). Socioeconomic Status and Health: What We Know and What We Don't. Annals of the New York Academy of Sciences, 896(1), 3–15. https://doi.org/10.1111/j.1749-6632.1999.tb08101.x

Brunekreef, B., & Holgate, S. T. (2002). Air pollution and health. The Lancet, 360(9341), 1233–1242. https://doi.org/10.1016/s0140-6736(02)11274-8

Finch, W., & Finch, M. (2019). Regularization Methods for Fitting Linear Models with Small Sample Sizes: Fitting the Lasso Estimator using R. Practical Assessment, Research, and Evaluation, 21(1). https://doi.org/10.7275/jr3d-cq04

Jha, P., Ramasundarahettige, C., Landsman, V., Rostron, B., Thun, M., Anderson, R. N., McAfee, T., & Peto, R. (2013). 21st-century hazards of smoking and benefits of cessation in the United States. The New England Journal of Medicine, 368(4), 341–350. https://doi.org/10.1056/NEJMsa1211128

Lantz, P. M., House, J. S., Lepkowski, J. M., Williams, D. R., Mero, R. P., & Chen, J. (1998). Socioeconomic Factors, Health Behaviors, and Mortality. JAMA, 279(21), 1703. https://doi.org/10.1001/jama.279.21.1703

McGinnis, J. M., Williams-Russo, P., & Knickman, J. R. (2002). The Case For More Active Policy Attention To Health Promotion. Health Affairs, 21(2), 78–93. https://doi.org/10.1377/hlthaff.21.2.78

Mokdad, A. H. (2004). Actual Causes of Death in the United States, 2000. JAMA, 291(10), 1238. https://doi.org/10.1001/jama.291.10.1238

Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., Mullany, E. C., Biryukov, S., Abbafati, C., Abera, S. F., Abraham, J. P., Abu-Rmeileh, N. M. E., Achoki, T., AlBuhairan, F. S., Alemu, Z. A., Alfonso, R., Ali, M. K., Ali, R., Guzman, N. A., & Ammar, W. (2014). Global, regional, and National Prevalence of Overweight and Obesity in Children and Adults during 1980-2013: a Systematic Analysis for the Global Burden of Disease Study 2013. Lancet (London, England), 384(9945), 766–781. https://doi.org/10.1016/S0140-6736(14)60460-8

Onwuegbuzie, A. J., & Daniel, L. G. (2019). Typology of Analytical and Interpretational Errors in Quantitative and Qualitative Educational Research. Current Issues in Education, 6. https://cie.asu.edu/ojs/index.php/cieatasu/article/view/1609

Pope III, C. A. (2002). Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution. JAMA, 287(9), 1132. https://doi.org/10.1001/jama.287.9.1132

Ross, C. E., & Mirowsky, J. (1995). Does Employment Affect Health? Journal of Health and Social Behavior, 36(3), 230. https://doi.org/10.2307/2137340

Shi, L., & Starfield, B. (2001). The Effect of Primary Care Physician Supply and Income Inequality on Mortality Among Blacks and Whites in US Metropolitan Areas. American Journal of Public Health, 91(8), 1246–1250. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1446754/

STIEFEL, M. C., PERLA, R. J., & ZELL, B. L. (2010). A Healthy Bottom Line: Healthy Life Expectancy as an Outcome Measure for Health Improvement Efforts. Milbank Quarterly, 88(1), 30–53. https://doi.org/10.1111/j.1468-0009.2010.00588.x

Walker, R. E., Keane, C. R., & Burke, J. G. (2010). Disparities and access to healthy food in the United States: A review of food deserts literature. Health & Place, 16(5), 876–884. https://doi.org/10.1016/j.healthplace.2010.04.013

Wilkinson, R., Marmot, M., & Europe, W. H. O. R. O. for. (2003). Social determinants of health: the solid facts. In iris.who.int. World Health Organization. Regional Office for Europe. https://iris.who.int/handle/10665/326568

World Health Organization. (2024). Social Determinants of Health. World Health Organisation. https://www.who.int/health-topics/social-determinants-of-health#tab=tab_1