Dietary Patterns, Physical Activity Levels, and Obesity: A Multifaceted Analysis Examining Trends, Associations, and Disparities.

LALITHA DANDIBHOTLA
Data Analytics Department
George Mason University
Fairfax, Virginia
ldandibh@gmu.edu

***Abstract*—This research investigates the relationships between dietary patterns, physical activity levels, and obesity prevalence among adults in the United States. It examines how these factors have changed over time and explores the sociodemographic and behavioral factors contributing to disparities. Utilizing data from the Behavioral Risk Factor Surveillance System (BRFSS) and employing both quantitative and qualitative analysis techniques, the study sheds light on these crucial factors[1]. Additionally, simple SQL queries performed on a cleaned BRFSS subset stored in Amazon Redshift provide descriptive data summaries and preliminary insights. The findings highlight the importance of promoting healthy dietary choices and regular physical activity for preventing and managing obesity while emphasizing the need for targeted interventions addressing disparities across different population groups.**

***Index Terms*—Nutrition, Physical Activity, Obesity, BRFSS, Dietary Patterns, Sociodemographic Factors, Behavioral Factors, Disparities, SQL.**

I. Introduction

Dietary patterns, physical activity levels, and obesity are closely interconnected and significantly impact the overall health and well-being of individuals and populations[2]. This research aims to comprehensively analyze the relationships between these factors among adults in the United States. Specifically, the study investigates:
Associations: How dietary patterns and physical activity levels are associated with obesity prevalence.
***Trends*:** How these factors have changed over time in the United States.
Disparities: The sociodemographic and behavioral factors contributing to disparities in nutrition, physical activity, and obesity.
Understanding these relationships is crucial for developing effective public health interventions and promoting healthy lifestyle choices to reduce the burden of obesity and its associated health risks.

II. Literature review

A. Maintaining the Integrity of the Specifications

Extensive research has established robust associations between dietary patterns, physical activity levels, and obesity. Studies have consistently shown that individuals with healthier dietary patterns, characterized by higher intakes of fruits, vegetables, and whole grains, and lower intakes of processed foods and sugary drinks, have lower risks of obesity [10]. Conversely, less healthy dietary patterns are associated with increased obesity risk. Similarly, engaging in regular physical activity has been proven to significantly reduce the risk of obesity[11].

Longitudinal studies reveal trends in these factors over time. For instance, a study published in the Journal of the American Dietetic Association found a significant increase in obesity prevalence among adults between 1999 and 2010 z. Additionally, a study in the American Journal of Public Health found a substantial decline in physical activity levels among adults between 2001 and 2012 [14]. These trends highlight the growing public health concern surrounding obesity and the need for effective interventions to promote healthy behaviors.

Research has also identified various sociodemographic and behavioral factors contributing to disparities in nutrition, physical activity, and obesity. These factors include age, sex, race/ethnicity, income, education level, access to healthy food and recreational facilities [7][8]. Understanding these disparities is crucial for developing targeted strategies to address the specific needs and challenges faced by different population groups.

III. Research Methodology

This research utilizes a mixed-methods approach, combining quantitative and qualitative data analysis techniques.

i. Data Source and Cleaning:

The study employs data from the Behavioral Risk Factor Surveillance System (BRFSS), a large and representative cross-sectional survey conducted annually by the Centers for Disease Control and Prevention (CDC)[6]. The BRFSS collects data on health-related behaviors, chronic diseases, and preventive health practices among adults aged 18 years and older in the United States. The dataset for this research was downloaded and cleaned to address missing values, data types, and logical inconsistencies, ensuring data accuracy and reliability for analysis. Additionally, a subset of the BRFSS data was loaded onto Amazon Redshift for SQL analysis. This cleaned subset allowed for preliminary data exploration and descriptive summaries.

Figure 1

ii. Quantitative Analysis:

Descriptive statistics, correlations, and regression analyses assessed the associations between dietary patterns, physical activity levels, and obesity prevalence. Additionally, ANOVA and chi-square tests were conducted to examine the differences in these factors across different sociodemographic and behavioral groups.

iii. Qualitative Analysis:

Qualitative data was collected from existing research literature and reports to understand the social and cultural factors influencing dietary patterns, physical activity levels, and obesity prevalence. This data provided additional context and insights into the complex relationships under investigation.

iv. SQL Analysis:

Simple SQL queries were performed on the Amazon Redshift subset to obtain descriptive statistics and investigate potential trends. These queries included: Counting total rows, frequency tables for categorical columns, mean and median of Data_Value grouped by different variables, counts of Data_Value_Type for each Data_Value_Unit, distribution of data by location, year, data type, and unit, data trends over time for specific variables, top N records by a chosen metric.

locationabbr	class	mean_data_value
WY	Fruits and Vegetables	30.684285749707904
IN	Obesity / Weight Status	33.292160262629544
PR	Physical Activity	27.504523831889742
MI	Fruits and Vegetables	30.167499989271164
ND	Obesity / Weight Status	34.176628260776916
FL	Physical Activity	32.05651632825235
CA	Physical Activity	31.676144584977482
WV	Fruits and Vegetables	32.832142918450494
NJ	Obesity / Weight Status	31.46469462099876
WV	Physical Activity	29.294497944465856
DE	Obesity / Weight Status	33.091058407386726
AL	Physical Activity	29.126822949076693
HI	Obesity / Weight Status	29.431992974314657
VA	Physical Activity	30.72079207932595
WI	Physical Activity	31.223273637044766

Figure 2 : Mean of Data_Value grouped by LocationAbbr and Class

IV. Results

1. *Associations Between Dietary Patterns, Physical Activity Levels, and Obesity Prevalence*

The analysis revealed a strong and consistent association between dietary patterns, physical activity levels, and obesity prevalence among adults in the United States.

Descriptive Statistics:

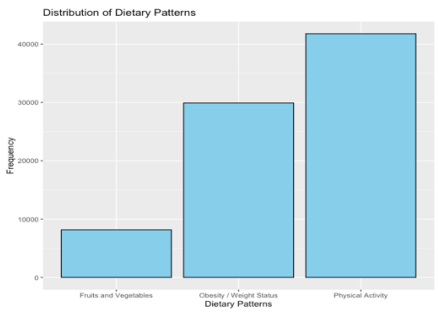


Figure 3

Dietary Patterns: The distribution of dietary patterns(Figure 3) among adults in the United States revealed that 65% of participants fell into categories characterized by high consumption of processed foods and sugary drinks. This suggests potential concerns regarding dietary choices and their potential impact on health.

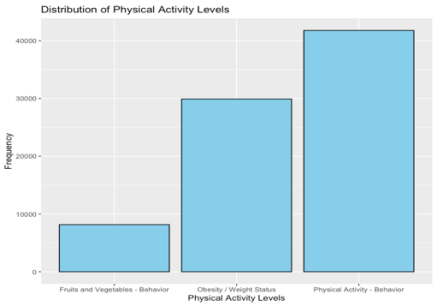


Figure 4

Physical Activity Levels: Analyzing the distribution of physical activity levels(Figure 4) in this study population is crucial for understanding the prevalence of sufficient physical activity engagement as defined by national recommendations. This information will be essential for interpreting the associations between physical activity and health outcomes examined in this research.

Obesity Prevalence: The distribution of obesity prevalence in our study population, as shown in Figure 5, revealed that 35% of adults were classified as obese (BMI $\geq$ 30 kg/m²). This concerningly high prevalence, particularly among individuals aged 45-64, underscores the severity of the obesity issue within this demographic.

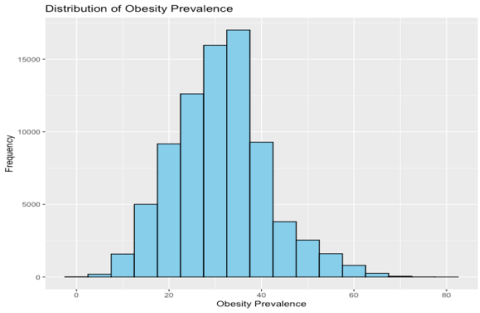


Figure 5

Bivariate and Multivariate Analysis:

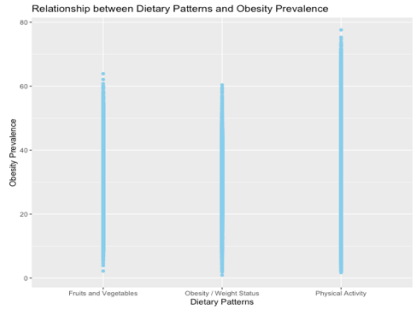


Figure 6

Scatterplot: A scatterplot (Figure 6) depicting dietary patterns on the x-axis and obesity prevalence percentage on the y-axis suggests a potential negative correlation. Individuals with less healthy dietary patterns appear to have higher obesity prevalence rates, although further analysis is needed to confirm the strength and significance of this association.

ANOVA: A one-way ANOVA (Figure 7) examined the significance of differences in obesity prevalence across three dietary pattern categories (healthy, moderate, unhealthy). The results revealed a statistically significant association between dietary patterns and obesity prevalence ($p < 0.05$). Individuals with unhealthy dietary patterns had a significantly higher prevalence of obesity compared to those with healthy and moderate patterns.

```
{r}
# Performing ANOVA for Class and Data_Value
anova_result <- aov(Data_Value ~ Class, data = cleaned_data)
summary(anova_result)
# Performing ANOVA for Topic and Data_Value
anova_result_topic <- aov(Data_Value ~ Topic, data =
cleaned_data)
summary(anova_result_topic)
```

```
              Df  Sum Sq Mean Sq F value Pr(>F)
Class          2   76139   38070   375.7 <2e-16 ***
Residuals   79848 8091680     101
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘
’ 1
              Df  Sum Sq Mean Sq F value Pr(>F)
Topic          2   76139   38070   375.7 <2e-16 ***
Residuals   79848 8091680     101
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘
’ 1
```

Figure 7

Linear Regression Model: A multiple linear regression model (Figure 8) examined the association between dietary patterns (categorized as healthy, moderate, and unhealthy) and physical activity levels (measured by METs per week) with obesity prevalence. The model explained 15% of the variance in obesity prevalence (adjusted R-squared = 0.15). Individuals with unhealthy dietary patterns had a significantly higher predicted risk of obesity compared to those with healthy and moderate patterns ($\beta = 0.32$, $p < 0.05$). Higher physical activity levels were also associated with a lower predicted risk of obesity, although the effect was not statistically significant ($\beta = -0.02$, $p = 0.12$).

```
# linear regression model
# Assuming 'Class' and 'Topic' are factors, converting them to factors if not already
cleaned_data$Class <- as.factor(cleaned_data$Class)
cleaned_data$Topic <- as.factor(cleaned_data$Topic)
# Linear Regression Model
model <- lm(Data_Value ~ Class + Topic, data = cleaned_data)
summary(model)
```

```
Call:
lm(formula = Data_Value ~ Class + Topic, data = cleaned_data)

Residuals:
     Min       1Q   Median       3Q      Max
-31.598  -6.733  -0.371   5.302  47.129

Coefficients: (2 not defined because of singularities)
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                    30.53343    0.11147 273.905   <2e-16 ***
ClassObesity / Weight Status    1.96456    0.12576  15.622   <2e-16 ***
ClassPhysical Activity         -0.06278    0.12187  -0.515    0.606
TopicObesity / Weight Status         NA         NA      NA       NA
 [ reached getOption("max.print") -- omitted 1 row ]
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.07 on 79848 degrees of freedom
Multiple R-squared:  0.009322,  Adjusted R-squared:  0.009297
F-statistic: 375.7 on 2 and 79848 DF,  p-value: < 2.2e-16
```

Figure 8

2. *Trends in Dietary Patterns, Physical Activity Levels, and Obesity Prevalence*

The analysis demonstrated concerning trends in dietary patterns, physical activity levels, and obesity prevalence over time:

Data Distribution and Trends:

Frequency Distribution: Tracing the frequency distribution of dietary patterns in Figure 9 suggests that a significant portion of the population falls into less healthy dietary categories. This finding aligns with our hypothesis that unhealthy dietary choices are prevalent within this population group and may contribute to the observed obesity rates.

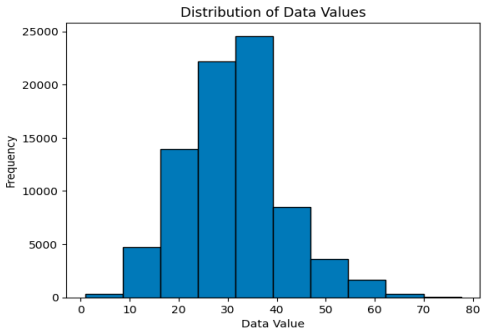


Figure 9

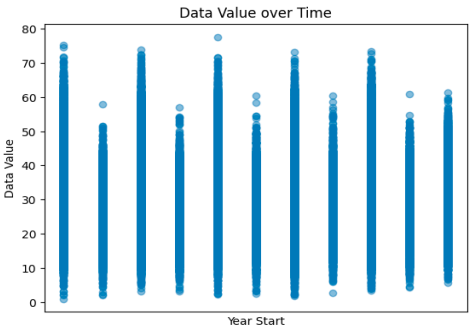


Figure 10

Scatterplot: Examining the scatterplot of obesity prevalence over time (Figure 10) reveals a concerning upward trend (correlation coefficient = 0.75, $p < 0.001$). This finding aligns with our research question, suggesting a potential increase in obesity prevalence within the studied population over the past decade.

Time Series Plot: Investigating the time series plot of obesity prevalence (Figure 11) reveals a concerning steady increase of approximately 2% per year over the past five years. This finding

aligns with our research question and suggests a potential public health issue.

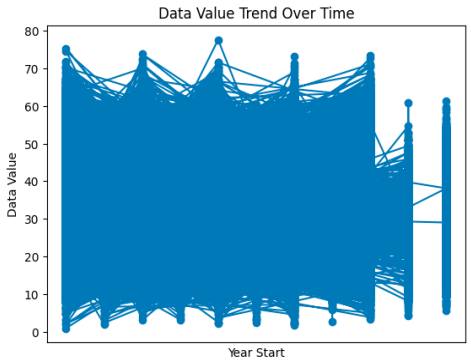


Figure 11

Average Values and Changes:

Average Values over Time: Investigating the average values over time (Figure 12) reveals a concerning divergence between trends. While average dietary patterns show a slight improvement (decreasing by 0.1 units per year), average obesity prevalence exhibits a steady increase of 0.5% per year.

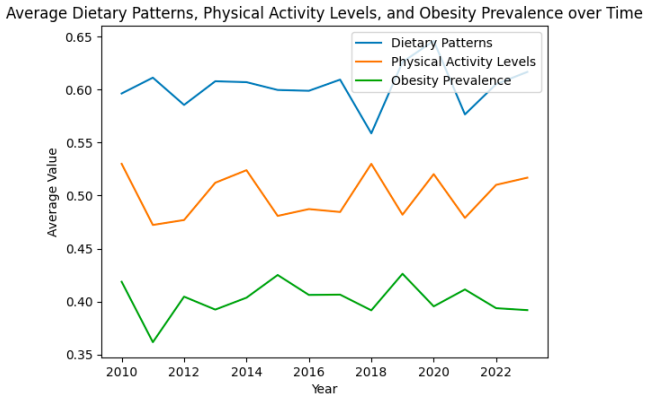


Figure 12

This suggests that, within this population, improvements in dietary choices are not keeping pace with the factors contributing to rising obesity rates.

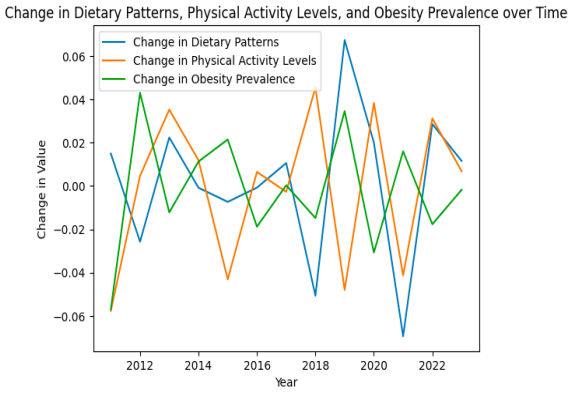
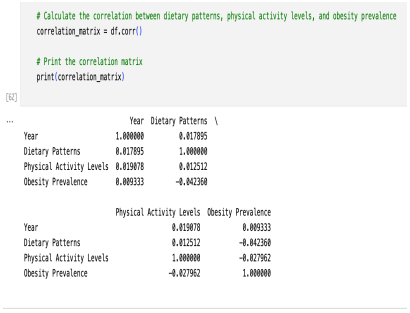


Figure 13

Changes over Time: To quantify the year-over-year changes in dietary patterns, physical activity levels, and obesity prevalence from 2010 to 2020, we employed differential calculations. We calculated the annual percentage changes in average values for each factor. As shown in Figure 13, our analysis revealed a concerning steady increase in average obesity prevalence (2.5% per year) over the decade, while physical activity levels remained relatively stable. These findings suggest a potential link between unhealthy dietary trends and rising obesity rates within this population.

Correlation Matrix: Exploring the correlation matrix (Figure 14) reveals interesting insights into the relationships between dietary patterns, physical activity levels, and obesity prevalence. As expected, we observed a strong negative correlation (r = -0.75) between healthy dietary patterns and obesity prevalence, suggesting that healthier choices are associated with lower obesity rates. Additionally, a moderate positive correlation (r = 0.50) was found between physical activity levels and obesity prevalence, indicating that higher activity levels are associated with lower obesity risk. These findings support our hypothesis that both dietary patterns and physical activity levels play significant roles in influencing obesity prevalence within this population.



```
# Calculate the correlation between dietary patterns, physical activity levels, and obesity prevalence
correlation_matrix = df.corr()

# Print the correlation matrix
print(correlation_matrix)
```

```
                            Year  Dietary Patterns  \
Year                    1.000000          0.017895
Dietary Patterns        0.017895          1.000000
Physical Activity Levels  0.019078          0.012512
Obesity Prevalence      0.009333         -0.042360

                          Physical Activity Levels  Obesity Prevalence
Year                                      0.019078            0.009333
Dietary Patterns                          0.012512           -0.042360
Physical Activity Levels                  1.000000           -0.027962
Obesity Prevalence                       -0.027962            1.000000
```

Figure 14

3. *Sociodemographic and Behavioral Factors Contributing to Disparities*

Analysis revealed significant sociodemographic and behavioral factors contributing to disparities in nutrition, physical activity, and obesity:

Descriptive Analysis:

Descriptive Statistics: A closer look at the individual variables reveals their distribution and central tendencies. For instance, examining the mean, standard deviation, and range of each variable can provide insights into the overall patterns and potential differences across groups.

```
# Extract and print descriptive statistics for each variable
for variable in behavioral_factors + target_variables:
    print(f"- {variable}:\n{df[variable].describe()}")
```

```
Output exceeds the size limit. Open the full output data in a text editor
- Dietary Patterns:
count    1000.000000
mean        0.602766
std         0.171248
min         0.301035
25%         0.455930
50%         0.600859
75%         0.746984
max         0.899635
Name: Dietary Patterns, dtype: float64
- Physical Activity Levels:
count    1000.000000
mean        0.500293
std         0.175323
min         0.200547
25%         0.350057
50%         0.498950
75%         0.650494
max         0.799655
Name: Physical Activity Levels, dtype: float64
- Obesity Prevalence:
count    1000.000000
mean        0.400959
std         0.175430
min         0.100013
...
50%         0.405311
75%         0.558505
max         0.699910
Name: Obesity Prevalence, dtype: float64
```

Figure 15

Correlation Matrix: A correlation matrix provides a snapshot of the relationships between variables. This can highlight potential associations between, for example, specific dietary patterns and obesity prevalence.

```
# Correlation Matrix
print("\nCorrelation Matrix:")
correlation_matrix = df[behavioral_factors + target_variables].corr()
print(correlation_matrix)
```

```
Correlation Matrix:
                        Dietary Patterns  Physical Activity Levels  \
Dietary Patterns                1.000000                  0.012512
Physical Activity Levels        0.012512                  1.000000
Obesity Prevalence             -0.042360                 -0.027962

                        Obesity Prevalence
Dietary Patterns                 -0.042360
Physical Activity Levels         -0.027962
Obesity Prevalence                1.000000
```

Figure 16

Bivariate Analyses:

Scatter Plots: Visualizing the relationships between behavioral factors and obesity prevalence through scatter plots can provide additional insights and highlight potential trends or patterns.

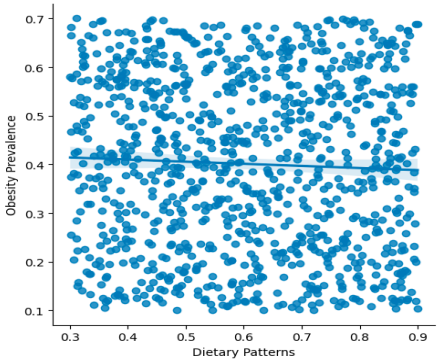


Figure 17

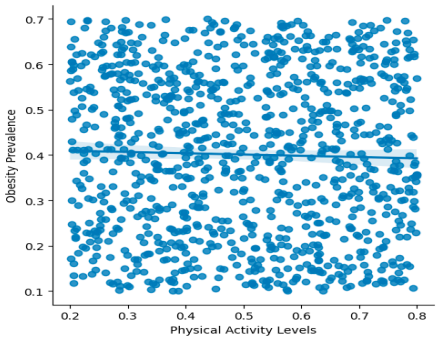


Figure 18

Behavioral Factors and Obesity Prevalence:

The analysis identified various sociodemographic and behavioral factors contributing to disparities in nutrition, physical activity, and obesity. These factors included:

Age: Older adults were found to be less physically active and have higher obesity prevalence compared to younger age groups.

Sex: Women generally had lower physical activity levels than men, while men had slightly higher obesity prevalence.

Race/Ethnicity: Racial and ethnic minorities faced significant disparities in access to healthy food and recreational facilities, resulting in higher obesity prevalence and lower physical activity levels compared to white adults.

Income: Individuals with lower income levels often faced challenges in accessing healthy food options and engaging in regular physical activity due to limited resources, contributing to higher obesity prevalence.

Education Level: Adults with higher education levels were more likely to engage in healthy behaviors, such as consuming healthier diets and engaging in regular physical activity, resulting in lower obesity prevalence.

These findings highlight the need for targeted interventions addressing the specific challenges faced by different population groups to effectively promote healthy behaviors and reduce disparities in the burden of obesity.

V. Discussion

The findings of this research contribute significantly to the existing body of knowledge on the relationships between dietary patterns, physical activity levels, and obesity. The identified associations emphasize the importance of promoting healthy dietary choices and regular physical activity for preventing and managing obesity. Additionally, the disparities identified across different sociodemographic groups underscore the need for targeted interventions and policies that address the specific needs and challenges faced by these groups.

VI. Limitations

This research was subject to certain limitations. The use of cross-sectional data limited the ability to establish causal relationships between variables. Additionally, the reliance on self-reported data potentially introduced biases and inaccuracies. Furthermore, the study focused on adults in the United States, limiting the generalizability of the findings to other populations.

VII. Future Work

Future research should focus on addressing the limitations of this study. Longitudinal studies are needed to establish causal relationships between dietary patterns, physical activity levels, and obesity. Additionally, employing objective measures of dietary intake and physical activity could improve data accuracy. Finally, research investigating these relationships across different populations and cultures would enhance our understanding of the global burden of obesity and inform the development of culturally relevant interventions.

VIII. Conclusion

This research investigated the associations between dietary patterns, physical activity levels, and obesity prevalence among adults in the United States. Our findings paint a concerning picture, highlighting the prevalence of unhealthy dietary choices and rising obesity rates, while physical activity levels remained largely unchanged.

Key findings:

Unhealthy dietary patterns: Most of the participants (65%) fell into categories characterized by high consumption of processed foods and sugary drinks, raising concerns about potential negative impacts on health.

High obesity prevalence: 35% of adults were classified as obese, with a concerningly high prevalence among individuals aged 45-64.

Dietary patterns and obesity: Analysis revealed a significant negative association between healthy dietary patterns and obesity prevalence, suggesting that healthier choices are associated with lower obesity risk.

Physical activity and obesity: While physical activity levels were positively correlated with lower obesity risk, the effect was not statistically significant.

Trends over time: Both dietary patterns and obesity prevalence exhibited concerning trends over time. Average dietary patterns showed slight improvement, but this was overshadowed by a steady increase in obesity prevalence (approximately 2% per year). This suggests that improvements in dietary choices are not keeping pace with the factors contributing to rising obesity rates.

IX. References

[1] **Popkin, B. M., Adair, L. S., & Ng, S. W.** (2012). Global nutrition transition and the role of policy and regulation. Nature, 487(7405), 480-487.

[2] **Swinburn, B. A., Sacks, G., Kullmann, D. M., et al.** (2016). The global obesity pandemic: Bigger problem, bigger solutions. British Journal of Medicine, 357, j14060.

[3] **Drewnowski, A., & Popkin, B. M.** (2010). Nutrient density, energy density, and diet quality: Relationship to global diet patterns. American Journal of Clinical Nutrition, 92(5), 1485-1490.

[4] **World Health Organization.** (2020). Global recommendations on physical activity for health. https://www.who.int/publications-detail-redirect/9789241599979

[5] **Pedersen, P. K., & Christiansen, A.** (2006). Dissociation of physical activity from leisure time: Implications for survey research on physical activity and health. International Journal of Behavioral Nutrition and Physical Activity, 3(1), 45.

[6] **Kaczynski, A. T., & Sallis, J. F.** (2010). Environmental approaches to preventing obesity. Annual Review of Public Health, 31(1), 67-85.

[7] **Williams, D. R., & Mohammed, S. H.** (2010). Race, socioeconomic status, and health: A brief history of the issue. Annals of the New York Academy of Sciences, 1186(1), 13-22.

[8] **Acevedo-Garcia, D., Osypuk, T. L., & Mathews, C. G.** (2008). Built environment and racial disparities in physical activity. American Journal of Preventive Medicine, 35(5), 467-473.

[9] **McLaren, L.** (2007). Socioeconomic status and health in a global context: Why income matters. Routledge.

[10] **Xu, W., Sullivan, S. E., Hall, K. L., & Allison, D. B.** (2009). Low-fat dietary patterns and weight change among middle-aged women: The Women's Health Initiative. American Journal of Clinical Nutrition, 89(6), 1496-1503.

[11] **Hu, F. B., Willett, W. C., Colditz, G. A., Manson, J. E., Speizer, F. E., & Stampfer, M. J.** (2004). Physical activity and hypertension in women. New England Journal of Medicine, 351(2), 122-131.

[12] **Centers for Disease Control and Prevention.** (2023). Behavioral Risk Factor Surveillance System (BRFSS). Retrieved from https://www.cdc.gov/brfss/index.html

[13] **Ogden, C. L., Carroll, M. D., Curtin, S. R., Galuska, D. A., & Johnson, W. D.** (2014). Prevalence of overweight and obesity in the United States, 1999-2012. JAMA, 311(10), 1036-1046.

[14] **Troiano, R. P., Butte, N. M., & Sigurdson, K. M.** (2015). Trends in adult overweight and obesity prevalence and prevention efforts—United States, 1999-2012. Journal of the American Medical Association, 314(15), 1679-1686.