Is High BMI in Children Aged 6-17 Years Associated with Poor Poverty Status?

By:
Resham Basnet

A Paper presented to the faculty of The University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the Degree of Master of Public Health in the Department of Maternal and Child Health

Chapel Hill, North Carolina

April 5th, 2018

Approved by:
Is High BMI in Children Aged 6-17 Years Associated with Poor Poverty Status?

Prepared by:
Resham Basnet
MPH Candidate, 2018
Department of Maternal and Child Health
Gillings School of Global Public Health
The University of North Carolina at Chapel Hill

Abstract

Objectives: The prevalence of childhood obesity and overweight has been increasing over the past two decades in the US, with 17.2% of children having high Body Mass Index (BMI) defined as 85% or higher in 2014. This paper assesses the association between poverty status and BMI in children aged 6-17 years.

Methods: The 2014 full year consolidated data from Medical Expenditure Panel Survey (MEPS) was analyzed to determine the association between poverty status and BMI using logistic regression while controlling for age, gender, and race/ethnicity variables. The analysis incorporated both sample weights and survey design adjustments.

Results: After holding all other variables in the model constant, children living below 100 percent poverty threshold had 1.9 times the odds of high BMI compared to children living at or above 400 percent poverty threshold. This association was statistically significant at p<0.01 [95% CI: 1.5, 2.2].

Conclusion: The results from logistic regression support the hypothesis that high BMI is prevalent among children with poor poverty status. This finding suggests that further research is needed to understand vulnerable populations while promoting healthy eating and physical activity.
ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere thanks to my academic advisor, Professor Kavita Singh Ongechi. Thank you for the continuous support, suggestions and encouragement throughout my MPH study and Master's paper. I have been extremely lucky to have you as my advisor who responded to my questions and queries and provided guidance during the course of my studies. I truly appreciate your mentorship.

I extend my deepest thanks to Professor Sally Clark Stearns. Thank you for introducing me to Regression models and teaching me how to perform regression analysis. I owe a debt of gratitude to you for your time and valuable instructions to finish this paper. Thank you for your support and guidance.

I would also like to thank my loving parents, Gehendra Bahadur Basnet and Shanta Devi Basnet, for their love and constant encouragement and motivation. My heartfelt thanks to my brother-in-law, Surendra Bahadur Kunwar, and sister-in-law, Bimita Thapa, for their love, support and encouragement.

Finally, my special thanks go to my wife, Manju Kunwar, and our daughter, Ramona Basnet, for their love, patience, and understanding. They are always by my side providing me lovingly support and make my life meaningful. They are the inspiration behind everything I do.
# Table of Contents

**ACRONYMS** ........................................................................................................................................... 6

**Introduction** ........................................................................................................................................ 7

**Methods** ............................................................................................................................................... 10

**Results** ............................................................................................................................................... 13

**Discussion and Conclusion** .............................................................................................................. 16

**References** ......................................................................................................................................... 19

**Tables:**

Table 1: Descriptive characteristics of children aged 6-17 in the US, stratified by body mass index ....... 21

Table 2: Weighted proportions stratified by BMI and Logistic regression analysis for high body mass index in children aged 6-17 years ................................................................................................. 22

Table 3: Comparison between Logistic Model and Logistic Model with Young Kids Interaction with Poverty status ............................................................................................................................................ 23

Table 4: Akaike's information criterion (AIC) and Bayesian information criterion (BIC) ................. 24

**Figures:**

Figure 1: Conceptual model on the relationship between poverty and childhood obesity .................. 25

Figure 2: Predictive Margins with 95% CI for the logistic regression model ................................................. 26

Figure 3: Predictive Margins with 95% CI for the logistic regression model with interaction between young kids of ages less than 12 and poverty status ................................................................................. 27

**Stata do file** ........................................................................................................................................ 28
LIST OF TABLES AND FIGURES

Tables:
- Table 1: Descriptive characteristics of children aged 6-17 in the US, stratified by body mass index
- Table 2: Weighted proportions stratified by BMI and Logistic regression analysis for high body mass index in children aged 6-17 years
- Table 3: Comparison between Logistic Model and Logistic Model with Young Kids Interaction with Poverty status
- Table 4: Akaike's information criterion (AIC) and Bayesian information criterion (BIC)

Figures:
- Figure 1: Conceptual model on the relationship between poverty and childhood obesity
- Figure 2: Predictive Margins with 95% CI for the logistic regression model
- Figure 3: Predictive Margins with 95% CI for the logistic regression model with interaction between young kids of ages less than 12 and poverty status
ACRONYMS:

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHQR</td>
<td>Agency for Healthcare Research and Quality</td>
</tr>
<tr>
<td>AIC</td>
<td>Akaike's Information Criterion</td>
</tr>
<tr>
<td>BIC</td>
<td>Bayesian Information Criterion</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>CBPR</td>
<td>Community-Based Participatory Research</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>MEPS</td>
<td>Medical Expenditure Panel Survey</td>
</tr>
<tr>
<td>MEPS HC</td>
<td>Medical Expenditure Panel Survey Household Component</td>
</tr>
<tr>
<td>SES</td>
<td>Socioeconomic Status</td>
</tr>
<tr>
<td>TEE</td>
<td>Total Energy Expenditure</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
</tbody>
</table>
Introduction

Childhood obesity is one of the major public health issues in the United States (US) that not only affects health of children as they grow but also is associated with higher healthcare expenditure.\textsuperscript{1,2} Body Mass Index (BMI), which is a screening tool that is often used to determine overweight and obesity, and is calculated as weight in kilograms divided by height in meters squared.\textsuperscript{3} Normal BMI in children falls between the 5\textsuperscript{th} percentile to the 85\textsuperscript{th} percentile, BMI for the overweight falls between the 85\textsuperscript{th} percentile to the 95\textsuperscript{th} percentile, and obese children have BMI equal to or greater than the 95\textsuperscript{th} percentile.\textsuperscript{4} In 2013-2014, 17.2\% of children in the US were found to have high BMI defined as BMI greater than 85\textsuperscript{th} percentile.\textsuperscript{2,4} Childhood obesity rates in the United States have more than tripled over the past three decades.\textsuperscript{5,6} The obesity rate among children aged 6-11 years increased four times, from 4.0\% in 1971-1974 to 18.8\% in 2003-2004, while the prevalence of obesity among children aged 12–19 years increased about three times, from 6.1\% in 1971-1974 to 17.4\% in 2003–2004.\textsuperscript{5}

Several factors have been demonstrated to play an important role with childhood obesity including socioeconomic status (SES), racial and ethnic background, age, gender, parents’ education, and family structure, among others. Household income plays a key factor in childhood obesity as it gives the ability for a family to purchase desired and healthy food. Studies have shown that low socioeconomic status groups are prone to develop high BMI in developed nations like the US.\textsuperscript{2,3} This inverse relationship between income and obesity is more prominent in developed countries, whereas a positive relationship between income and obesity is usually seen among individuals in developing nations. Change in diet with a rise in consumption of foods rich in fats and sugar such as fast foods, snacks, and sugary beverages by children is believed to be one of the factors associated with childhood obesity and excess weight gain.\textsuperscript{7} Children with high
BMI are more likely to grow obese as adults and are at higher risk of developing various health conditions such as heart disease, hypertension, type 2 diabetes, stroke and different types of cancer during childhood and as they age. It is therefore important to have a good understanding of the association between poverty status and BMI, because knowledge of the association may help target public health interventions and development of health policy.

In the conceptual model, as shown in figure 1, poverty directly affects the level of education and attitude towards healthy diet. Similarly, poverty also affects getting healthcare access and health literacy about healthy diet and exercise. Likewise, race and ethnicity also affect the poverty level and the choices of food. All these factors affect the type of food choices and the amount of physical activity performed among children. Unhealthy food consumption and limited physical activity, in turn, affect child’s weight and BMI. The other control variables ‘age’ and ‘gender’ are also used in this study. It is evident that biological differences between males and females are apparent in the development of fat patterning. Gender difference is seen in many areas of nutrition in childhood and adolescence. High fat-free mass of males is associated with high energy intake and expenditure. Different bodies of males and females may impact the ability to engage in certain physical or sporting activities. A longitudinal study found that daily total energy expenditure (TEE) for males increased continuously between ages 5 and 10, but, for females, TEE increased from around 1,400 kcal at age 5 to 1,800 kcal at 6, but by age 9, it decreased to 1,600 kcal. This was explained by a 50% reduction in physical activity between ages 6 and 9. It has also been suggested that the level of physical activities in male kids of ages 11 to 12 are nearly twice that of their female counterparts. Studies have also found that females are more likely to pay attention to food calories, whereas males eat more fast food. Age and
gender differences have also been noted in body image perception, weight concern and weight control practices among teenagers.\textsuperscript{10}

Social inequalities in obesity and overweight prevalence has substantially increased in the US. One study found that the obesity prevalence was increased from 23\% in 2003 to 33\% in 2007 for children in low-education, low-income, and higher unemployment households, and overweight prevalence was increased from 13\% in 2003 to 15\% in 2007 for children in low-education and low-income households.\textsuperscript{11} It also found that the prevalence of childhood obesity was 30.4\% in 2007 among children with parents having less than 12 years of education, 3.1 times higher than the prevalence among children with parents having a college degree.\textsuperscript{11} Similarly, the prevalence of childhood obesity among children below 100\% poverty threshold was 27.4\%, 2.7 times higher than the prevalence among children with poverty level greater than 400\% of the poverty threshold.\textsuperscript{11} Additionally, the study also found that almost half of all children in low-education and low-income group were overweight, compared with less than 23\% of children in high-education or high-income group.\textsuperscript{11}

Social inequality in childhood obesity is not just limited to education and income levels. The inequality has also been observed across different races and ethnicity. In 2007, the prevalence of childhood obesity among Asian children was found to be a low of 8.7\% whereas the prevalence was higher for Black and Hispanic children of 23.9\% and 23\%, respectively.\textsuperscript{11} Similarly, childhood overweight prevalence varied from 18.4\% for Asian children to 41\% for Black and Hispanic children. Between 2003 and 2007, the prevalence of childhood obesity among Hispanic children has swelled by 24.2\%.\textsuperscript{11} This increase in prevalence of childhood obesity over last few decades clearly shows that the magnitude of racial and ethnic disparities in childhood obesity and overweight prevalence has increased markedly over the time.
The purpose of this paper is to examine the relationship between poor poverty status and childhood obesity and to provide estimates and changes in obesity and overweight prevalence among children aged 6-17 years using the nationally representative samples of US children in 2014. By performing a secondary data analysis from the Medical Expenditure Panel Survey Household Component (MEPS HC) data for the year 2014, I hypothesize that children of ages 6-17 living below 100 percent poverty threshold have increased prevalence of BMI greater than 85th percentile when compared to children living at or above 400 percent threshold in the US, after controlling for variables including age, gender, race/ethnicity, and health literacy about diet and exercise.

**Methods**

To determine the association between poverty status and high BMI in children aged 6-17 years, a secondary analysis of the household component from 2014 Full Year Consolidated Data File: MEPS HC-171 dataset was performed. The Household Component of the 2014 MEPS, a survey sponsored by the Agency for Healthcare Research and Quality (AHQR), is a nationally representative sample of the US civilian noninstitutionalized population. A total of 13,421 household (33,162 respondents) were interviewed in 2014, and data were collected in five rounds of interview over the two year of survey participation. The MEPS-HC provides information on health status, demographic and socio-economic characteristics, health services utilization, and health care spending and uses an overlapping panel design to collect data from a nationally representative sample of households.

All the variables from the 2014 Full Year Consolidated Data File: MEPS HC-171 dataset were reviewed, and the variables that were most relevant to this analysis were included in the
The primary dependent and independent variables included in this study are variable describing BMI in children aged 6-17 years and poverty status. In this study, the dependent variable ‘childhood obesity’ is defined as BMI greater than the 85th percentile which is a combined BMI for both overweight and obese and the independent variable ‘poor poverty status’ is based on the ratio of income to poverty and is below 100% poverty threshold. Furthermore, other variables including age, gender, race/ethnicity, and health literacy about diet and exercise were included in this model as covariates because these variables are found to have an association between poverty status and BMI.13

After cleaning the dataset, variables were re-named, re-coded, and re-constructed based on the model for this study. The main dependent variable for this analysis is BMI in children aged 6-17 years, which was calculated as weight in kilograms divided by height in centimeters divided by 100 squared. BMI for children of 6-17 years of age was already calculated in MEPS based upon parentally reported weight and height, which was further dichotomized (85th percentile or less and more than 85th percentile) based on the recommended Centers for Disease Control and Prevention (CDC) growth chart for boys and girls.4 According to the CDC growth chart, BMI in children is based on the sex and age, and any BMI over 85th percentiles fall into overweight criteria.4 Using this information, high BMI variable was dichotomized as ‘yes’ (>85th percentile) and ‘no’ (≤85th percentile) for both boys and girls. The main independent variable for this study is poverty status which is a categorical variable with five categories based on the ratio of income to poverty: poor (below 100% of poverty), near poor (100% to 125% of poverty), low income (125% to 200% of poverty), middle income (200% to 400% of poverty), and high income (at least 400% of poverty).
The survey included information about doctor or other health provider giving advice about eating healthy and the amount and kind of exercise, sports or physically active hobbies the child should have. These variables were dichotomized as ‘yes’ (advice given) and ‘no’ (advice not given). Education was categorized into three age groups: elementary school age group (ages 6 to 8 years), middle school age group (9 to 12 years), and high school age group (13 to 17 years). Gender was dichotomized into males and females. Lastly, racial and ethnicity was grouped into five categories (Hispanic, non-Hispanic White only, non-Hispanic Black only, non-Hispanic Asian only, and non-Hispanic other race or multi-race).

MEPS-HC 2014 dataset had a total of 34,875 observations; children aged 6-17 had 6,712 observations. Around 25% (1,686) of the BMI variable had missing height and/or weight information to calculate BMI, and therefore was not included in the complete case analysis. A total of 4,964 observations were used for the complete case analysis. All statistical analyses were performed using STATA version 14.1. Descriptive statistics (frequencies and percentages) are provided for variables by the main outcome, as shown on Table 1. Additionally, Pearson’s chi-square test was used to calculate differences in variables by body mass index status and are provided with p-values (Table 1). Logistic regression analyses were performed to test for the association between independent variables and the outcome of interest. Odds ratios, 95% confidence intervals, and p-values were calculated and reported, as shown on Table 2. A multiple logistic regression analysis was conducted controlling for all variables. Odds ratios, 95% confidence intervals, and p-values were calculated for the final multivariate model and reported in Table 2.

I hypothesized that children aged 6-17 years with ‘poor’ poverty status (living below 100 percent of poverty threshold) have high BMI (greater than 85th percentile) compared to children
living at or above 400 percent poverty threshold. To determine this association, I used two models using Stata: logistic model, and logistic model with interaction between young kids of ages less than 12 years and ‘poor’ poverty status (below 100 percent poverty threshold). The interaction was done after generating a new variable of young kids that included children of ages less than 12. As BMI can vary in different age groups, this interaction will help in understanding the predicted probabilities of age (young kids of age<12 years) on the poor poverty status (below 100 percent of poverty threshold) and high BMI (greater than 85th percentile). Since MEPS uses complex survey design, I incorporate both sample weight and survey design. The Akaike's information criterion (AIC) and Bayesian information criterion (BIC) was used to choose the preferred final regression model. Then, I ran F-test hypothesis to determine whether the association was statistically significant and different from the null value. Subsequently, marginal effect and 95% confidence interval was used to predict the association. Finally, I predicted the relationship between poverty status and BMI in children aged 6-17 years.

Results

Table 1 summarizes descriptive characteristics of children aged 6-17 years, stratified by body mass index. Of the 6,712 children aged 6-17 years, 33% (n=2,219) reported to being under 100 percent of poverty threshold and about two-third have BMI greater than 85th percentile. For race/ethnicity, the majority of children was Hispanic (37%) followed by non-Hispanic Whites only (28%), non-Hispanic Blacks only (23%), non-Hispanic multi-race (7%), and non-Hispanic Asian only (5%). The prevalence of high BMI varied from a low of 3.6% in non-Hispanic Asian children to a high of 45.5% in Hispanic children. It is also noted in Table 1 that the prevalence of BMI is comparatively higher among children in elementary and middle school (6-12 years old) than to children in high school (13-17 years old).
Table 2 shows logistic regression analysis for high BMI in children aged 6-17 years, which includes the number of non-missing observations, values of weighted proportions stratified by BMI, and bivariate association with the p-value from the F-test. The weighted proportion of high BMI in children aged 6-17 years was 36.2%, 7.3%, 20.5%, 22.9%, and 13.1% among poor, near poor, low income, middle income, and high income respectively. Children below 100 percent of poverty account for 36% of children with high BMI but only 26% of children with low BMI. In contrast, children at or above 400 percent of poverty account for only 13% of children with high BMI while more than 20% of children with low BMI. The weighted proportion of high BMI was 18.9%, 35.3%, and 45.8% for children in elementary school, middle school, and high school respectively. Additionally, Hispanic children account for 41% of children with high BMI but only 34% of children with low BMI. Similarly, Non-Hispanic Black children account for 25% of children with high BMI but only 20% of children with low BMI. In contrast, Non-Hispanic Asian children account for only 3% of children with high BMI but more than 7% of children with low BMI.

The results from the logistic regression adjusted for weight and survey design, and the adjusted Wald test (F-test) was used to assess the association between the main dependent variable, the main independent variable, and other covariates, as shown in Table 2. The association between high BMI and poverty status indicated that after holding all other variables in the model constant, children living below 100 percent poverty threshold had 1.9 times the odds of high BMI compared to children living at or above 400 percent poverty threshold. This association was statistically significant at OR 1.85 [95% CI: 1.52, 2.24] p<0.01. Likewise, the odds of high BMI when compared with Hispanic children was comparatively lower for non-Hispanic Asian and non-Hispanic White children. The odds of high BMI are 0.4 times for non-
Hispanic Asian children (OR 0.40; 95% CI: 0.29, 0.5; p<0.001) and 0.79 times for non-Hispanic White children (OR 0.79; 95% CI: 0.67, 0.92; p<0.001) when compared to Hispanic children.

The adjusted Wald test assessing the categorical poverty status variable showed that we can reject the null hypothesis that the poverty status coefficients are jointly equal to zero in favor of the alternative that the coefficients of these variables are not all equal to zero (F-value = 11.8, p-value <0.0001).

The findings from the two models: logistic model, and logistic with interaction between young kids (children less than 12 years of age) and poverty status are presented in Table 3. Both models consistently showed children living below 100 percent poverty threshold are more likely to have BMI greater than 85th percentile when compared to those living at or above 400 percent poverty threshold. After holding all other variables in the logistic model constant, children living below 100 percent poverty level had 1.85 times the odds of having BMI greater than 85th percentile when compared to those living at or above 400 percent poverty level. Similarly, after holding all other variables in the logistic model with interaction between young kids of ages less than 12 years and poverty status constant, children living below 100 percent poverty level had 1.65 times the odds of having BMI greater than 85th percentile when compared to children living at or above 400 percent poverty level. Both associations are statistically significant at p<0.01 [OR 1.85; 95% CI: 1.52, 2.24] and [OR 1.65; 95% CI: 1.28, 2.12], respectively. Also, the interaction between young kids and poverty showed that young kids had 1.5 times the odds of high BMI compared to old kids. All findings were statistically significant at p<0.05.

Based on the findings from the AIC/BIC criteria, as shown in Table 4, the logistic model with interaction had lower values of AIC, and hence, was used as a preferred model. The adjusted Wald test assessing the categorical poverty status variable, for the logistic with
interaction model, showed that we can reject the null hypothesis that the poverty status coefficients are jointly equal to zero in favor of the alternative that the coefficients of these variables are not all equal to zero (F-value = 4.2, p-value = 0.002).

The Predictive margins and 95% CI, as shown in Figures 2 and 3, are a type of direct standardization in which the predicted values from the logistic regression model and logistic regression model with interaction between young kids and poverty represents the average predicted response of BMI over the distribution of different poverty status from ‘poor’ to ‘high income’. In the predictive margin graph from the logistic regression model, as shown in Figure 2, we see that the BMI is high in children with ‘poor’ poverty status and the BMI declines as the poverty status changes from ‘poor’ poverty status to ‘high income’ status. The predictive margin graph from the logistic regression model with interaction between young kids of ages <12 and poverty status also showed similar findings for both young kids of ages <12 and for kids of ages 12 and older. The graph predicted that the BMI declines in both age groups as the poverty status shifts from poverty status of below 100 percent threshold to at least 400 percent threshold.

**Discussion and Conclusion**

I hypothesized that children living below 100 percent of poverty have high prevalence of having BMI greater than 85\textsuperscript{th} percentile. Using 2014 MEPS-HC dataset, a secondary data analysis was performed to examine the association between poverty status and BMI for children in the US. The findings from this study supports my hypothesis that high BMI is more prevalent among children aged 6-17 years with poor poverty status compared to those with high income. These findings are also consistent with other US studies that compared the relationship between overweight and obesity in children and teens with socioeconomic status between 1971 and
This inverse relationship between high BMI and ‘poor’ poverty status was also observed in the predictive margin graph for both logistic regression models. Additionally, findings from this study also showed that disparities in high BMI exist for a specific racial/ethnic group; odds of having BMI greater than 85th percentile was higher among Hispanic children compared to non-Hispanic White and Asian children.

This study has few limitations. One of the limitations is that the recorded children’s height and/or weight data in this study were based on parents’ reports and were not independently measured. However, previous studies have shown that parental reported height and/or weight data can be used as a valid indicator to calculate BMI. Another limitation of this study is that around 25% of BMI observations were missing from the dataset, so these observations could not be included in the complete case analysis. The missing data results in a smaller sample size and could introduce bias unless the data are missing completely at random. Additionally, because of lack of data in the 2014 MEPS-HC, I was unable to consider some of the variables such as children’s dietary habits and choices (consumption of vegetables, fruits, and/or sugary foods/beverages), neighborhood socioeconomic conditions (access to park, playground, sidewalk, recreation center), and children’s physical activity and television viewing habits, all of which could potentially affect children’s high BMI risks.

In summary, this paper affirmed my hypothesis that high BMI is prevalent among children with poor poverty status, and differences among various race/ethnicity groups exist in the US. High BMI in children can affect their physical, social, and emotional well-being both during their childhood and as they continue to grow, reducing the life expectancy. High BMI is also responsible for higher healthcare expenditure. Considerable efforts and infrastructure are needed at the national, state and local levels to focus on implementing programs that promote
healthy eating and physical activity. Findings suggest that more research is required and focus should be geared towards the vulnerable population groups such as children with poor poverty status and among Hispanic population. Community-based participatory research (CBPR) has been shown to be effective in addressing socioeconomic and racial/ethnic disparities for various health conditions including childhood obesity at the individual and community level determinants.\textsuperscript{19-21} CBPR provides a mechanism for using the knowledge of the local health context, basic conceptual models, and research knowledge and helps in designing interventions that can be sustainable and cost-effective for bringing changes in communities.\textsuperscript{22} CBPR connects the knowledge of expert researchers with the experience of community leaders and partners, which can be an ideal approach in developing childhood obesity prevention intervention and thus increasing awareness and bridging the racial and income disparity gap across the US.
References


Table 1: Descriptive characteristics of children aged 6-17 in the US, stratified by body mass index (n=6,712) (MEPS-HC 2014)

<table>
<thead>
<tr>
<th></th>
<th>Body Mass Index</th>
<th></th>
<th></th>
<th></th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;85th percentile</td>
<td>&gt;85th percentile</td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poverty Status</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Below 100 percent of poverty (Poor)</td>
<td>793</td>
<td>35.7</td>
<td>1,426</td>
<td>64.3</td>
<td>2,219</td>
</tr>
<tr>
<td>100 to 125 percent of poverty (Near Poor)</td>
<td>219</td>
<td>43.4</td>
<td>285</td>
<td>56.6</td>
<td>504</td>
</tr>
<tr>
<td>125 to 200 percent of poverty (Low Income)</td>
<td>565</td>
<td>42.3</td>
<td>772</td>
<td>57.7</td>
<td>1,337</td>
</tr>
<tr>
<td>200 to 400 percent of poverty (Middle Income)</td>
<td>839</td>
<td>51.4</td>
<td>793</td>
<td>48.6</td>
<td>1,632</td>
</tr>
<tr>
<td>At least 400 percent of poverty (High Income)</td>
<td>625</td>
<td>61.3</td>
<td>395</td>
<td>38.7</td>
<td>1,020</td>
</tr>
<tr>
<td>Gender</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1,494</td>
<td>43.3</td>
<td>1,953</td>
<td>56.7</td>
<td>3,447</td>
</tr>
<tr>
<td>Female</td>
<td>1,547</td>
<td>47.4</td>
<td>1,718</td>
<td>52.6</td>
<td>3,265</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>1,012</td>
<td>37.8</td>
<td>1,669</td>
<td>62.2</td>
<td>2,681</td>
</tr>
<tr>
<td>Non-Hispanic White only</td>
<td>954</td>
<td>53.0</td>
<td>846</td>
<td>47.0</td>
<td>1,800</td>
</tr>
<tr>
<td>Non-Hispanic Black only</td>
<td>620</td>
<td>42.8</td>
<td>830</td>
<td>57.2</td>
<td>1,450</td>
</tr>
<tr>
<td>Non-Hispanic Asian only</td>
<td>230</td>
<td>63.7</td>
<td>131</td>
<td>36.3</td>
<td>361</td>
</tr>
<tr>
<td>Non-Hispanic other race or multi-race</td>
<td>225</td>
<td>53.6</td>
<td>195</td>
<td>46.4</td>
<td>420</td>
</tr>
<tr>
<td>Advice about healthy diet from provider</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advice received</td>
<td>1,834</td>
<td>44.8</td>
<td>2,261</td>
<td>55.2</td>
<td>4,095</td>
</tr>
<tr>
<td>Advice not received</td>
<td>1,177</td>
<td>48.3</td>
<td>1,260</td>
<td>51.7</td>
<td>2,437</td>
</tr>
<tr>
<td>Advice about physical activity from provider</td>
<td>0.169</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advice received</td>
<td>1,410</td>
<td>45.2</td>
<td>1,707</td>
<td>54.8</td>
<td>3,117</td>
</tr>
<tr>
<td>Advice not received</td>
<td>1,600</td>
<td>46.9</td>
<td>1,809</td>
<td>53.1</td>
<td>3,409</td>
</tr>
<tr>
<td>Education (Age)</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary school age group (6 to 8 years old)</td>
<td>586</td>
<td>33.3</td>
<td>1,173</td>
<td>66.7</td>
<td>1,759</td>
</tr>
<tr>
<td>Middle school age group (9 to 12 years old)</td>
<td>940</td>
<td>42.2</td>
<td>1,286</td>
<td>57.8</td>
<td>2,226</td>
</tr>
<tr>
<td>High school age group (13 to 17 years old)</td>
<td>1,515</td>
<td>55.6</td>
<td>1,212</td>
<td>44.4</td>
<td>2,727</td>
</tr>
</tbody>
</table>

*p-value from the Pearson’s chi-squared test
Table 2: Weighted proportions stratified by BMI and Logistic regression analysis for high body mass index in children aged 6-17 years (n=4,964) (MEPS HC 2014)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of non-missing observations</th>
<th>Weighted proportions</th>
<th>Logistic Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&lt;85&lt;sup&gt;th&lt;/sup&gt; percentile</td>
<td>&gt;85&lt;sup&gt;th&lt;/sup&gt; percentile</td>
</tr>
<tr>
<td><strong>Body Mass Index</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Poverty Status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below 100 percent of poverty (Poor)</td>
<td>2,219</td>
<td>25.9%</td>
<td>36.2%</td>
</tr>
<tr>
<td>100 to 125 percent of poverty (Near Poor)</td>
<td>504</td>
<td>7.3%</td>
<td>7.3%</td>
</tr>
<tr>
<td>125 to 200 percent of poverty (Low Income)</td>
<td>1,337</td>
<td>18.8%</td>
<td>20.5%</td>
</tr>
<tr>
<td>200 to 400 percent of poverty (Middle Income)</td>
<td>1,632</td>
<td>27.7%</td>
<td>22.9%</td>
</tr>
<tr>
<td>At least 400 percent of poverty (High Income)</td>
<td>1,020</td>
<td>20.3%</td>
<td>13.1%</td>
</tr>
<tr>
<td><strong>Education (Age)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary school age group (6 to 8 years old)</td>
<td>1,759</td>
<td>15.8%</td>
<td>18.9%</td>
</tr>
<tr>
<td>Middle school age group (9 to 12 years old)</td>
<td>2,226</td>
<td>28.8%</td>
<td>35.3%</td>
</tr>
<tr>
<td>High school age group (13 to 17 years old)</td>
<td>2,727</td>
<td>55.4%</td>
<td>45.8%</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3,447</td>
<td>48.7%</td>
<td>53%</td>
</tr>
<tr>
<td>Female</td>
<td>3,265</td>
<td>51.3%</td>
<td>47%</td>
</tr>
<tr>
<td><strong>Health literacy about diet</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advice received</td>
<td>4,095</td>
<td>60.5%</td>
<td>68.5%</td>
</tr>
<tr>
<td>Advice not received</td>
<td>2,437</td>
<td>39.5%</td>
<td>31.5%</td>
</tr>
<tr>
<td><strong>Health literacy about physical activity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advice received</td>
<td>3,117</td>
<td>46.8%</td>
<td>53.3%</td>
</tr>
<tr>
<td>Advice not received</td>
<td>3,409</td>
<td>53.2%</td>
<td>46.7%</td>
</tr>
<tr>
<td><strong>Race/Ethnicity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>2,681</td>
<td>33.7%</td>
<td>40.6%</td>
</tr>
<tr>
<td>Non-Hispanic White only</td>
<td>1,800</td>
<td>30.8%</td>
<td>25.2%</td>
</tr>
<tr>
<td>Non-Hispanic Black only</td>
<td>1,450</td>
<td>20.8%</td>
<td>25.2%</td>
</tr>
<tr>
<td>Non-Hispanic Asian only</td>
<td>361</td>
<td>7.4%</td>
<td>3.1%</td>
</tr>
<tr>
<td>Non-Hispanic Other race or multi-race</td>
<td>420</td>
<td>7.3%</td>
<td>5.9%</td>
</tr>
</tbody>
</table>
### Table 3: Comparison between Logistic Model and Logistic Model with Interaction between Young Kids ages <12 years and Poverty status (n=4,964) (MEPS HC 2014)

<table>
<thead>
<tr>
<th>High BMI</th>
<th>Logistic Model OR [95% CI]</th>
<th>Logistic Model with Interaction OR [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young kids ages &lt;12 years</td>
<td>1.53** [1.14, 2.05]</td>
<td></td>
</tr>
<tr>
<td>Poverty Status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below 100 percent of poverty (Poor)</td>
<td>1.85*** [1.52, 2.24]</td>
<td>1.65*** [1.28, 2.12]</td>
</tr>
<tr>
<td>100 to 125 percent of poverty (Near Poor)</td>
<td>1.39* [1.05, 1.82]</td>
<td>1.21 [0.84, 1.73]</td>
</tr>
<tr>
<td>125 to 200 percent of poverty (Low Income)</td>
<td>1.53*** [1.24, 1.87]</td>
<td>1.50*** [1.15, 1.95]</td>
</tr>
<tr>
<td>200 to 400 percent of poverty (Middle Income)</td>
<td>1.19* [0.98, 1.44]</td>
<td>1.19 [0.92, 1.53]</td>
</tr>
<tr>
<td>At least 400 percent of poverty (High Income)</td>
<td>Reference</td>
<td>Reference</td>
</tr>
</tbody>
</table>

**Interaction between Young kids ages <12 years and poverty status**

<table>
<thead>
<tr>
<th></th>
<th>Logistic Model OR [95% CI]</th>
<th>Logistic Model with Interaction OR [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 100 percent of poverty (Poor)</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>100 to 125 percent of poverty (Near Poor)</td>
<td>1.07 [0.66, 1.75]</td>
<td></td>
</tr>
<tr>
<td>125 to 200 percent of poverty (Low Income)</td>
<td>0.81 [0.58, 1.13]</td>
<td></td>
</tr>
<tr>
<td>200 to 400 percent of poverty (Middle Income)</td>
<td>0.76 [0.56, 1.05]</td>
<td></td>
</tr>
<tr>
<td>At least 400 percent of poverty (High Income)</td>
<td>0.76 [0.53, 1.09]</td>
<td></td>
</tr>
</tbody>
</table>

**Gender (male)**

<table>
<thead>
<tr>
<th></th>
<th>Logistic Model OR [95% CI]</th>
<th>Logistic Model with Interaction OR [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.22*** [1.08, 1.37]</td>
<td>1.22*** [1.08, 1.37]</td>
<td></td>
</tr>
</tbody>
</table>

**Race/Ethnicity**

<table>
<thead>
<tr>
<th></th>
<th>Logistic Model OR [95% CI]</th>
<th>Logistic Model with Interaction OR [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hispanic</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Non-Hispanic White only</td>
<td>0.79*** [0.67, 0.92]</td>
<td>0.79*** [0.67, 0.92]</td>
</tr>
<tr>
<td>Non-Hispanic Black only</td>
<td>0.99 [0.84, 1.15]</td>
<td>0.98 [0.84, 1.15]</td>
</tr>
<tr>
<td>Non-Hispanic Asian only</td>
<td>0.40*** [0.29, 0.54]</td>
<td>0.40*** [0.29, 0.54]</td>
</tr>
<tr>
<td>Non-Hispanic other race or multi-race</td>
<td>0.70** [0.54, 0.90]</td>
<td>0.70** [0.54, 0.90]</td>
</tr>
</tbody>
</table>

**Education (Age)**

<table>
<thead>
<tr>
<th></th>
<th>Logistic Model OR [95% CI]</th>
<th>Logistic Model with Interaction OR [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary school age group (6 to 8 years old)</td>
<td>Reference</td>
<td></td>
</tr>
<tr>
<td>Middle school age group (9 to 12 years old)</td>
<td>1.18 [0.93, 1.49]</td>
<td></td>
</tr>
<tr>
<td>High school age group (13 to 17 years old)</td>
<td>0.97 [0.64, 1.47]</td>
<td></td>
</tr>
</tbody>
</table>

**Health literacy about diet**

<table>
<thead>
<tr>
<th></th>
<th>Logistic Model OR [95% CI]</th>
<th>Logistic Model with Interaction OR [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.35*** [1.15, 1.58]</td>
<td>1.36*** [1.16, 1.59]</td>
<td></td>
</tr>
</tbody>
</table>

**Health literacy about physical activity**

<table>
<thead>
<tr>
<th></th>
<th>Logistic Model OR [95% CI]</th>
<th>Logistic Model with Interaction OR [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.14* [0.98, 1.32]</td>
<td>1.14* [0.98, 1.32]</td>
<td></td>
</tr>
</tbody>
</table>

*** p<0.01, ** p<0.05, * p<0.1
Table 4: Akaike's information criterion (AIC) and Bayesian information criterion (BIC)

<table>
<thead>
<tr>
<th>Model</th>
<th>Observations</th>
<th>ll (null)</th>
<th>ll(model)</th>
<th>df</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Model</td>
<td>4,964</td>
<td>-72902.34</td>
<td>-70325.26</td>
<td>15</td>
<td>140680.5</td>
<td>140778.2</td>
</tr>
<tr>
<td>Logistic Model with Interaction</td>
<td>4,964</td>
<td>-72902.34</td>
<td>-70242.94</td>
<td>20</td>
<td>140525.9</td>
<td>140656.1</td>
</tr>
</tbody>
</table>
Figure 1: Conceptual model on the relationship between poverty and childhood obesity.
Figure 2: Predictive Margins with 95% CI for the logistic regression model
Figure 3: Predictive Margins with 95% CI for the logistic regression model with interaction between young kids of ages less than 12 and poverty status.
Stata do file

keep chbmix42 povcat14 age14x sex racethx eathlt42 physcl42

**Now, I want to generate variables relevant to my study population (which is children aged 6-17)

gen bmi=chbmix42 if age14x>=6 & age14x<=17
gen poverty=povcat14 if age14x>=6 & age14x<=17
gen age=age14x if age14x>=6 & age14x<=17
gen gender=sex if age14x>=6 & age14x<=17
gen race=racethx if age14x>=6 & age14x<=17
gen diet=eathlt42 if age14x>=6 & age14x<=17
gen exercise=physcl42 if age14x>=6 & age14x<=17

**I will now replace missing values for bmi (-1 and -9, which are classified in MEPS as either inapplicable and missing height or weight to calculate BMI, respectively), so they are "numerically missing" in stata.

replace bmi=. if bmi== -1 | bmi== -9
replace diet=. if diet== -1 | diet== -9 | diet== -8 | diet== -7
replace exercise=. if exercise== -1 | exercise== -9 | exercise== -8 | exercise== -7

**Now, I will recode gender (dichotomous) variable into 0/1 dummies

recode gender (1=1 "Male") (2=0 "Female"), gen(male)
tab male

**Recode other variables
recode poverty (1=1 "Poor") (2=2 "Near Poor") (3=3 "Low Income") (4=4 "Middle Income") (5=5 "High Income"),gen(catpoverty)

tab catpoverty

recode exercise (1=1 Yes) (2=0 No),gen(adviceexcercise)

tab adviceexcercise

recode diet (1=1 Yes) (2=0 No),gen(advisediet)

tab advisediet

recode age (6/8=1 "Elementary school age") (9/12=2 "Middle school age") (13/17=3 "High school age"),gen(catage)

tab catage

#delimit;

*generate high_bmi for girls;

gen high_bmi=.;
replace high_bmi=1 if male==0&age==6&bmi>=17.1;
replace high_bmi=1 if male==0&age==7&bmi>=17.6;
replace high_bmi=1 if male==0&age==8&bmi>=18.3;
replace high_bmi=1 if male==0&age==9&bmi>=19.1;
replace high_bmi=1 if male==0&age==10&bmi>=20;
replace high_bmi=1 if male==0&age==11&bmi>=20.9;
replace high_bmi=1 if male==0&age==12&bmi>=21.7;
replace high_bmi=1 if male==0&age==13&bmi>=22.6;
replace high_bmi=1 if male==0&age==14&bmi>=23.4;
replace high_bmi=1 if male==0&age==15&bmi>=24.0;
replace high_bmi=1 if male==0&age==16&bmi>=24.7;
replace high_bmi=1 if male==0&age==17&bmi>=25.2;
replace high_bmi=0 if male==0&age==6&bmi<17.1;
replace high_bmi=0 if male==0&age==7&bmi<17.6;
replace high_bmi=0 if male==0&age==8&bmi<18.3;
replace high_bmi=0 if male==0&age==9&bmi<19.1;
replace high_bmi=0 if male==0&age==10&bmi<20;
replace high_bmi=0 if male==0&age==11&bmi<20.9;
replace high_bmi=0 if male==0&age==12&bmi<21.7;
replace high_bmi=0 if male==0&age==13&bmi<22.6;
replace high_bmi=0 if male==0&age==14&bmi<23.4;
replace high_bmi=0 if male==0&age==15&bmi<24.0;
replace high_bmi=0 if male==0&age==16&bmi<24.7;
replace high_bmi=0 if male==0&age==17&bmi<25.2;

*for boys:

replace high_bmi=1 if male==1&age==6&bmi>=17.0;
replace high_bmi=1 if male==1&age==7&bmi>=17.4;
replace high_bmi=1 if male==1&age==8&bmi>=17.9;
replace high_bmi=1 if male==1&age==9&bmi>=18.6;
replace high_bmi=1 if male==1&age==10&bmi>=19.4;
replace high_bmi=1 if male==1&age==11&bmi>=20.2;
replace high_bmi=1 if male==1&age==12&bmi>=21.0;
replace high_bmi=1 if male==1&age==13&bmi>=21.9;
replace high_bmi=1 if male==1&age==14&bmi>=22.7;
replace high_bmi=1 if male==1&age==15&bmi>=23.4;
replace high_bmi=1 if male==1&age==16&bmi>=24.2;
replace high_bmi=1 if male==1&age==17&bmi>=24.9;

replace high_bmi=0 if male==1&age==6&bmi<17.0;
replace high_bmi=0 if male==1&age==7&bmi<17.4;
replace high_bmi=0 if male==1&age==8&bmi<17.9;
replace high_bmi=0 if male==1&age==9&bmi<18.6;
replace high_bmi=0 if male==1&age==10&bmi<19.4;
replace high_bmi=0 if male==1&age==11&bmi<20.2;
replace high_bmi=0 if male==1&age==12&bmi<21.0;
replace high_bmi=0 if male==1&age==13&bmi<21.9;
replace high_bmi=0 if male==1&age==14&bmi<22.7;
replace high_bmi=0 if male==1&age==15&bmi<23.4;
replace high_bmi=0 if male==1&age==16&bmi<24.2;
replace high_bmi=0 if male==1&age==17&bmi<24.9;

replace high_bmi=. if bmi==.

**Running test for table 1 (descriptive)
tab catpoverty high_bmi, row chi2
    tab male high_bmi, row chi2
tab race high_bmi, row chi2
tab advicediet high_bmi, row chi2
tab adviceexcercise high_bmi, row chi2
tab catage high_bmi, row chi2

    gen mysample_cca=e(sample)==1

    replace mysample_cca=age~=
    tab mysample
** Summarizing all variables to find out unweighted means and proportions for column 2

```
sum high_bmi catpoverty catage race male advicediet adviceexcercise if mysample_cca==1

tab catpoverty high_bmi if mysample_cca==1, row chi2

tab race high_bmi if mysample_cca==1, row chi2

tab male high_bmi if mysample_cca==1, row chi2

tab high_bmi if mysample_cca==1

tab advicediet high_bmi if mysample_cca==1, row chi2

tab adviceexcercise high_bmi if mysample_cca==1, row chi2

tab catage high_bmi if mysample_cca==1, row chi2
```

**Now, I am calculating weighted means and proportions for column 3. Since, I want to look at BMI, I used bmi to weight.

```
svyset [pweight=bmi]

svy, subpop(age): mean high_bmi catpoverty age14x race male advicediet adviceexcercise

svy, subpop(age): tab catpoverty high_bmi, col

svy, subpop(age): tab race high_bmi, col
```
svy, subpop(age): tab male high_bmi, col

svy, subpop(age): tab advicediet high_bmi, col

svy, subpop(age): tab adviceexcercise high_bmi, col

svy, subpop(age): tab catage high_bmi, col

**Running logistic regression test

svy, subpop(mysample): logistic high_bmi ib5.catpoverty i.male i.race age14x i.advicediet
   i.adviceexcercise i.catage

test 2.catpoverty 3.catpoverty 4.catpoverty 5.catpoverty

test 2.race 3.race 4.race 5.race

test 1.male

test 1.advicediet

test 1.adviceexcercise

test 2.catage 3.catage

*Logistic model
svy, subpop(mysample): logistic high_bmi ib5.catpoverty i.male i.race age14x i.advicediet i.adviceexcercise i.catage

outreg2 using paper_output, append ctitle("Logistic") eform

margins, over(catpoverty)
marginsplot

test 1.catpoverty 2.catpoverty 3.catpoverty 4.catpoverty 5.catpoverty

*generating young kids aged less than 12
gen youngkids=age<12
replace youngkids=. if age==.

*Logistic model w/ interaction

svy, subpop(mysample): logistic high_bmi ib5.catpoverty##i.youngkids ///
i.male i.race i.advicediet i.adviceexcercise age14x

outreg2 using paper_output, append ctitle("Logistic with interaction") eform
margins catpoverty#youngkids, subpop(mysample)
marginsplot

test 1.catpoverty 2.catpoverty 3.catpoverty 4.catpoverty 5.catpoverty

*AIC & BIC

quietly logistic high_bmi i.catpoverty i.male i.race age14x i.advicediet i.adviceexcercise i.catage
if mysample==1 [pweight=bmi]
estimates store model_logistic

quietly logistic high_bmi i.youngkids##i.catpoverty i.male i.race age14x i.advicexdiet
i.adviceexcercise i.catage if mysample==1  [pweight=bmi]
estimates store model_logisticwinteraction

estimates stats model_logistic model_logisticwinteraction