Predicting Obesity among Adolescents in the United States Using Modified Logistic Model

Eleanor K. Jator[*]

Austin Peay State University, Clarksville, TN 37043
*Corresponding author: jatore@apsu.edu

Abstract Obesity among adolescents is still on the rise and various reasons have been attributed to this increase. Obesity has been associated with many diseases, as well as, increase in healthcare costs. Concentration index and logistic regression have been extensively used to measure inequalities in health, including obesity, but these methods require each parameter to be calculated discretely. In this study, the logistic regression model is modified to predict the degree of obesity distribution that might be associated with multiple variables including income and race among adolescents in the United States. Unlike the methods currently used, the modified logistic model (MLM) can capture all variables at the same time in a single equation. The results produced by the model are comparable with those obtained when concentration index is used in a shorter time. It is hoped that this study will shorten the time to estimate or predict obesity rates among various races using existing Medical Expenditure Panel Survey (MEPS) data based on socioeconomic status. The ultimate goal is to develop targeted intervention strategies. Using existing data yields faster, reliable results since the sampling and collection utilize standard procedures. Results can easily be generalized due to random sampling and the MLM has the potential to predict the rate of obesity without performing further statistical analysis.

Keywords: modified logistic regression, Income, obesity, concentration index, socioeconomic inequality, adolescents

Cite This Article: Eleanor K. Jator, "Predicting Obesity among Adolescents in the United States Using Modified Logistic Model." *American Journal of Public Health Research*, vol. 2, no. 3 (2014): 86-90. doi: 10.12691/ajphr-2-3-4.

1. Introduction

Obesity, which is a risk factor for many diseases, is on the rise in the United States. The prevalence of obesity among adolescents in the United States remains high [1]. Globally, Obesity is associated with an increased risk of mortality and low life expectancy [2]. Obesity is known to occur when there is an energy imbalance. In which case, a positive energy imbalance occurs when the energy intake surpasses expenditure, resulting in weight gain [3]. According to the 1999-2008 National Health and Nutrition Examination Survey (NHANES), using heights and weights, eighteen percent of adolescents (12-19 years) are overweight in the United States [1].

Research studies have indicated that obese adolescents are at risk of becoming obese adults [4]. This implies that diseases occurring during childhood will likely persist in adulthood. Socioeconomic factors such as low income, low education, and type of occupation have been positively linked to obesity. Many studies have explored the role of socioeconomic status on obesity, and have positively correlated low income and low education to increase in adult body weight [5].

The association between poverty and poor health status had been well documented for many years [6]. The consequences of obesity in any age group are enormous. If this problem is not adequately addressed, the rate of diseases associated with obesity will increase, leading to higher medical bills and low self-esteem. Sorlie, Backlind and Keller [7] examined the trend of obesity related diseases and the associated economic impact on youths from ages 6 to 17 using the National Hospital Discharge Survey data for 1979 to 1999. Diseases examined in the study included diabetes, sleep apnea, gall bladder disease, heart disease and other obesity related diseases.

The results of the study indicated that the percentage of hospital releases due to obesity related diseases soared from 1979 to 1981 and from 1997 to 1999. The number of discharges due to diabetes rose twofold compared to the years before the obesity epidemic. The rate of obesity and gall bladder diseases increased approximately threefold, while sleep apnea increased fivefold. The obesity epidemic has enormous impact, not only on the individuals concerned, but also on communities.

The prevalence of obesity-related conditions emphasizes the need for aggressiveness in preventing and treating obesity instead of concentrating on the associated diseases. Obesity seems to be prevalent in certain socioeconomic classes and low in others.

The distribution of income has a profound effect on body weight and other morbidities. Socioeconomic inequalities studies have reported that communities with

huge socioeconomic variations have higher rates of mortality and morbidity compared to communities with less disparities in socioeconomic status. Individuals from the latter areas are more likely to experience lower morbidity and mortality rates [8].

The studies on adolescent obesity have used logistic regression as the statistical method in determining the relationship between income and obesity. However, logistic regression [9] is based on a predictive model that establishes a relationship between a dichotomous variable and other variables, but falls short to quantitatively determine income inequality in obesity distribution. Hence, the use of concentration index and Modified Logistic Model (MLM) are vital for quantifying the distribution of the disparities.

This study will examine the use of concentration index and a MLM to determine inequalities in obesity across the socioeconomic spectrum and race. Concentration index is an index that reflects inequalities in health based on socioeconomic status [10]. On the other hand, the Modified Logistic Model is constructed as a function of several predictors that can be used to measure the obesity distribution among different races based on socioeconomic status. It is anticipated that this study would broaden the understanding of the impact of socioeconomic inequality on adolescent obesity so that obesity preventive measures might be more targeted to the vulnerable population. Many intervention programs both at the national, community, and school levels have been implemented, but the trends in obesity still remain high. Unfortunately, many weight management programs are nonspecific. These programs are developed for all social classes; "one size fits all."

2. Materials and Methods

This study utilized existing data to classify obese adolescents into categories based on family income among different races. This study's data was obtained from the 2001 consolidated Medical Expenditure Panel Survey (MEPS) Household Component (HC). The Medical Expenditure Panel Survey (MEPS) is a dataset that provides valuable information on health care costs and use in the United States.

MEPS consists of four components: Household component (HC), Nursing Home Component, Medical Provider Component, and Insurance Component. The MEPS-HC contains data collected from some families and individuals from a nationally representative sub-sample of households that participated in the prior year's National Health Interview Survey (NHIS). The 2001 MEPS Household dataset consists of 33,556 records. The participants include children, adolescents, and adults. This study extracted data on body mass index (BMI), family income, race/ethnicity on all adolescents between the ages 12-17. There were 3121 records on adolescents aged 12-17 years but the study extracted only the adolescent population having complete information. Thus, 2911 adolescents were included in the study. Of this total, 1462 are females and 1449 are males.

This dataset is for public use at no cost and no permission is required to use the data. It is released as SPSS and SAS software compatible transport dataset. It can be retrieved from MEPS website at http://www.meps.ahrq.gov/PufFiles/H60/H60doc.htm.

Table 1. Code and Description of Variables Used in the Analysis

Variable	Code
Subject Number	**PID**
Obesity (0 = Body Mass Index (BMI) < 95 percentile, 1 = BMI ≥ 95 percentile)	*OBESE*
Age of Teenager (Years)	*AGE01*
Gender (1=Male, 2=Female)	*SEX*
Race (1=American Indian, 2=Aleut, Eskimo, 3=Asian or Pacific Islander, 4=Black and 5=White)	*RACEX*
Family Income as a Percent of Poverty Line (1=Poor/Negative, 2=Near Poor, 3=Low Income, 4=Middle Income, 5=High Income)	*POVCAT01*

SAS statistical software (SAS Institute Inc, Cary, NC) was used to extract the required records from the public used file. The variables used as shown in Table 1 included: Age (AGE01X), Race (RACEX), Gender (SEX), obesity (OBESE) and Family Income as a Percent of Poverty (POVCAT01).

Four variables (AGE01X, RACEX, SEX and POVCAT01) are possible predictors of obesity in teenagers (ages 12 to 17 inclusive) dichotomized as follows: 0 = Body Mass Index (BMI) < 95 percentile and 1 = BMI ≥ 95 percentile as described in the Center for Disease Control and Prevention (Center for Disease Control and Prevention, nd). Obesity is the only dependent variable, and was based on BMI values. BMI, which is the recommended measure to determine obesity and overweight in children and adolescents, is an indirect measure of adiposity using height and weight [weight (kg)/height (m^2)]. Information on BMI for children and adolescents has already been calculated from weight and height by MEPS. Obesity was defined as adolescents having a BMI of at least the 95th percentile for age and sex as defined by the Center for Disease Control and Prevention [11].

The main independent variables of this study are income and race. Household income was categorized by dividing family income by applicable poverty line. The resulting percentages were grouped into 5 categories including: poor (less than 100%), near poor (100% to less than 125%), low income (125% to less than 200%), middle income (200% to less than 400%), and high income (greater than or equal to 400%). Races used in the study include: Whites, Blacks, Asian or Pacific Islanders, Aleut or Eskimos.

Concentration Index

Measuring health inequalities represents the first step towards the identification of inequities in health. Studies on inequality have used a wide variety of measures for determining the magnitude of inequalities in health [7]. Concentration index (CI) has been reported to be valuable in measuring inequalities in health. CI portrays the socioeconomic aspect to inequalities in health and can be manipulated to reflect changes in the distribution of the

population across socioeconomic groups [12]. On the other hand, MLM can be used to obtain similar results in a shorter time with multiple predictor indicators' input.

Modified Logistic Model described

The model is given by

$$P = \frac{e^L}{1+e^L} \qquad (1)$$

Where P is the measure of obesity and L is a function of ten variables whose coefficients are given by

$$L = -1.7109 - 0.1300x1 - 0.3157x2 - 0.4000x3$$
$$- 0.6463x4 - 0.1285x5 - 0.4166x6$$
$$+ 0.6299x7 + 2.0936x8 - 0.4221x9 + 0.5413x10.$$

The predictors x1, x2, x3, and x4 represent family income as a percent of poverty Line (x1 = near poor, x2=low income, x3 = middle income, x4 = high income) and the predictors x7, x8, x9, and x10 represent race (x7=American Indian or Aleut, Eskimo, x8=Asian or Pacific Islander, x9=Black, x10=White). The additional predictors x5 and x6 represent age of teenagers (x5=Years) and gender (x6 = male or female).

This MLM is more flexible compared to the logistic regression model which is based on a binary outcome. This model has the ability to predict the degree of obesity given the appropriate predictors such as race, income level, gender and age at the same time. The slopes for calculating these attributes were obtained by performing statistical analysis using logistic regression.

3. Results

Income inequality in the distribution of obesity was examined using the concentration index and the MLM. The slopes of the models are given in Table 2.

Table 2. Logistic Regression: Estimated Slopes, Standard Errors, Odds Ratios, with OBESE as Dependent Variable

Covariate	Estimated Slope	Std Error of Slope	Odds Ratio (OR)	95% CI for OR		-2 Log-Likelihood	LR Test Statistics	P-Value
				Lower	Upper			
INTERCEPT	-1.7109	0.00058	N/A	N/A	N/A	19662450	N/A	N/A
SEX	-0.4166	0.00118	0.659	0.658	0.661	19535927	126523.06	**<0.0001**
RACE2[a]	0.6299	0.00496	1.877	1.859	1.896	19478243	184206.96	**<0.0001**
RACE3	2.0936	0.01810	8.114	7.831	8.407			
RACE4	-0.4221	0.00339	0.656	0.651	0.660			
RACE5	0.5413	0.00144	1.718	1.713	1.723			
POVERTY2[b]	-0.1300	0.00298	0.878	0.873	0.883	19523189	139261.36	**<0.0001**
POVERTY3	-0.3157	0.00202	0.729	0.726	0.732			
POVERTY4	-0.4000	0.00174	0.670	0.668	0.673			
POVERTY5	-0.6463	0.00180	0.524	0.522	0.526			
Age	-0.1285	0.00035	0.879	0.879	0.880	19521529	140921.00	**<0.0001**

The results in Table 2 were used to calculate the results given in Table 3 and Table 4. The concentration index (CI) of obesity is presented in Table 3. The maximum value the concentration index can take is +1 and this happens when the wealthiest group in the population experience poor health, while the minimum value the index can take is -1. The minimum index of -1 occurs when poor health is concentrated among the lower income individuals.

Table 3. Concentration Index of Obese Adolescents in the United States

Income Levels	Percent	Cumulative	Number of obese	Percent of obese	Cumulative	Concentration index
Poor/ Negative	12.81	12.81	620690	17.65	17.65	.0007
Near poor	4.14	16.95	180948	5.15	22.8	-.0074
Low income	15.54	32.49	582200	16.56	39.36	-.0241
Middle income	33.93	66.42	1184473	33.69	73.05	-.0663
High income	33.58	100	947564	26.95	100	0.0000
Total						-0.0985

Table 4. Predicted values of Obese Adolescents using the MLM (1)

Income Levels	Probability	Percent
Poor/ Negative	0.1370	13.70
Near poor	0.1223	12.23
Low income	0.1090	10.90
Middle income	0.0970	9.70
High income	0.0862	8.62

The CI among adolescents, the study sample, was -0.0985 (See Table 1), which was significant at p<0.05. This negative number implies that obesity was concentrated among adolescents from lower income bracket households. The concentration index is represented graphically in Figure 1, and in this case, the concentration curve lies above the diagonal since the number is negative, implying that adolescents from high income households are less at risk of becoming obese.

The obesity concentration curve is shown in Figure 1 with the curve lying above the diagonal. This curve is above the diagonal, which implies that obesity is concentrated among the poor. This negative index signifies that income inequality is detrimental to adolescents from poor, near poor, and low income households; indicating that adolescents from high income households are favored.
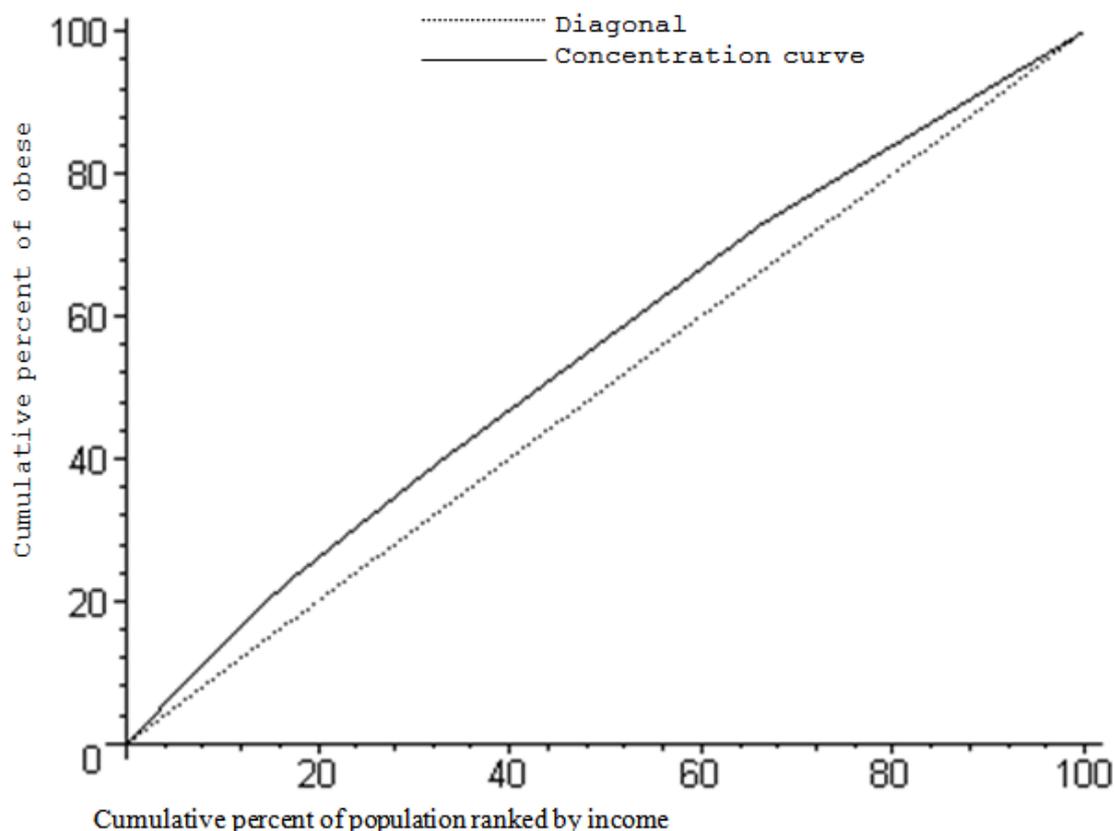
Figure 1. Obesity Concentration Curve in Adolescents

Table 5. Predicted values of Obese Adolescents (female, age:12 years)

Race	Probability	Percent
American Indian	**0.0679**	6.79
Aleut, Eskimo	**0.0679**	6.79
Asian or Pacific Islander	**0.9130**	91.30
Black	**0.2274**	22.74
White	**0.0036**	0.36

4. Discussion

The results produced by the MLM model are comparable with those obtained when concentration index is used, but MLM results are obtained in a shorter time. Furthermore, MLM has the ability to predict the degree of obesity given the appropriate predictors such as race, income level, gender and age at the same time as demonstrated in Table 4 and Table 5. It is worthwhile to note that the results given by the MLM are consistent with those given by the concentration index in Table 3. In either case, it was observed that there is an inverse association between income and obesity in adolescents. This finding confirms similar findings by other researchers [13] who reported that adolescents from low income households are more likely to be obese compared to those from high income households. Prevention and treatment of obesity is not only a medical problem but should be considered a public policy issue [14]. Targeting this obesity epidemic should be given a national priority followed by adequate funding of obesity reduction and prevention programs.

5. Conclusion

Although the concentration index and the MLM have been used to determine the degree of obesity inequality in adolescents, the MLM is more flexible in including multiple variables as predictors at the same time; while the concentration index is used to determine the disparities in obesity by quantifying obesity distribution using only a single variable as a predictor. The findings from the study can provide policy makers with vital information on the degree of socioeconomic inequality impact in obesity in this study group using MLM. Therefore, it is necessary to come up with improved methods, such as the MLM for capturing inequalities in obesity in order to develop programs that will be easily accessible and available to vulnerable individuals. The MLM can also be extended to study obesity disparities in adults.

Acknowledgements

Statement of Competing Interests

The author has no competing interests.

References

[1] Ogden CL, Carroll MD, Curtin LR, Lamb MM. & Flegal KM.(2010) Prevalence of high body mass index in U.S. children and adolescents, 2007-2008. JAMA, 303(3), 242-249.

[2] Raj, M. & Kumar, K. R. (2010). Obesity in children & Adolescents. *Indian J Med Res, 132*, 598-607.

[3] Flegal, K. M. (1999). The obesity epidemic in children and adults: current evidence and research issues. *Medicine Science Sports Exercise,* 31(11 Supplment 1).

[4] Sheperd, T. M. (2003). Effective management of obesity - Research findings that are changing clinical practice. Journal of family practice, 52(1), 34-42.

[5] Sundquist, J., & Johansson, S. (1998). The influence of socioeconomic status, ethnicity, and lifestyle on body mass index in a longitudinal study. *International journal of epidemiology, 27*, 57-63.

[6] Sorlie, P. D., Backlind, E. & Keller, J. B. (1995). U.S. mortality by economic, demographic, and social characteristics: The National Longitudinal Mortality Study. *American Journal of Public Health, 85*(7), 949-956.

[7] Wang, G. and. Dietz, W. H. (2002). Economic Burden of Obesity in Youths Aged 6 to 17 Years: 1979-1999. *Pediatrics, 109*(5), e81. Retrieved September 28, 2013 from http://pediatrics.aappublications.org/content/109/5/e81.full?sid=da 557998-4f1d-4fca-9236-360af4437c94.

[8] Kaplan, G, A., Pamuk, E, R., Lynch, J. W., Cohen, R. D. & Balfour, J. L. (1996). Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *British Medical Journal, 312*(7037), 999-1003.

[9] Moore, S. D. & McCabe, G. P. (2003). *Introduction to the Practice of Statistics.* W. H. Freeman and Company, New York.

[10] Wagstaff, A., Paci, P. & Doorslayer, E. V. (1991). On the measurement of inequalities in health. *Social Science & Medicine, 33*(5), 545-557.

[11] Kuczmarski, R. J., Ogben, C. L. & Guo, S. S. (2000). *CDC growth charts for the United States methods and development.* National Center for Health Statistics. Vital Health Stat 11. 2002; No 246.

[12] Zhang, Q. & Wang, Y. (2004). Socioeconomic inequality of obesity in the United States: do gender, age, and ethnicity matter? *Social Science & Medicine, 58*(6), 1171-1180.

[13] Goodman, E. (1999). The role of socioeconomic status gradient in explaining differences in US adolescent health (1999). *American Journal of Public health, 89*(10), 1522-1528.

[14] Williams, D. R. (1992). Black-white differentials in blood pressure: The role of social factors. *Ethnicity and Disease, 2*, 126-141.