Investigating Risk Factors Associated with Obesity Rate in the United States Based on PLS Regression and ANN Models

Tifini L. Batts, Chau-Kuang Chen, Juanita Buford, Vanisha L. Brown, and John Hughes, Jr.

Meharry Medical College,
Nashville, Tennessee, USA

ABSTRACT

This study was designed to assess potential risk factors for obesity in the United States via Partial Least Squares (PLS) Regression and Artificial Neural Network (ANN) Models. These factors included unemployment rate, poverty rate, individuals with less than a ninth grade education, those that lack a high school diploma, those that have a high school diploma, alcohol consumption per capita, percent of smokers in state, participation in physical exercise, violence/crime rate, average allotment of food stamps per month, and cost of hospital and physician services, independently, with respect to calendar year. Data was collected at the state level by federal agencies from years 2000 to 2006 (n = 350). Unemployment rate, poverty rate, alcohol consumption per capita, percent smokers in state, average allotment of food stamps per month, and cost of hospital and physician services with increase in time were statistically significant, had at least 50 percent of relative importance in relation to obesity rate, and may be used to predict obesity rate's severity. All significant risk factors, except alcohol consumption per capita, showed a positive association with obesity rate. This study negated claims of a significant relationship between the obesity rate and risk factors such as educational attainment, physical exercise, and the presence of violence/crime. It also proposed a positive effect from smoking and a negative effect from alcohol consumption per capita, whose direction of significance varied in previous studies.

INTRODUCTION

Obesity is a disease resulting from excess body fat (U.S. National Library of Medicine, 2013). It has been clinically identified according to body mass index (BMI), a ratio of an individual's weight to height, indirectly indicative of body fat composition (Weaver, Hayes, Unwin, & Murtagh, 2008). The qualifying BMI for obesity is a measure of greater than or equal to 30 kg/m^2 (Medical Encyclopedia, 2010).

Chronic ingestion of excess calories without adequate caloric usage may result in obesity. Besides diet, studies have shown that incident cases were a consequence of behavior, lack of exercise, and technological advancement, which promotes a more sedentary lifestyle (Sallis & Glanz, 2009; U. S. National Library of Medicine, 2013; World Health Organization [WHO], 2013). The occurrence of obesity also surges due to genetic predisposition (U.S. National Library of Medicine, 2013). Irrespective of cause, within the last couple of decades, obesity's occurrence has been steadily rising in the United States, corresponding to increased health complications and financial burden for those affected individuals (Centers

for Disease Control and Prevention [CDC], 2012; Notaro, et al., 2011; Pearson, Bhat-Schelbert, Ford, & Mokdad, 2009; Tsugane, 2012; WHO, 2013).

Problem Statement

In the United States, greater than one-third of adults have been categorized as obese, with the most dramatic addition occurring within the last 13 years (CDC, 2012). Its occurrence has been thought to disclose chronic conditions like heart disease, cancers, and stroke (CDC, 2012; Dixon, Dixon, & O'Brien, 2003; Mokdad, Ford, Bowman, Dietz, Vinicor, et al., 2003; Prospective Studies Collaboration, 2009; U.S. National Library of Medicine, 2013). As anticipated, the coexistence of obesity and its associated morbid diseases have led to decreased life expectancy, quality of life, and financial difficulties (Mokdad, Ford, Bowman, Dietz, Vinicor, et al., 2003; Peeters, Barendregt, Willekens, Mackenbach, Al Mamun, et al., 2003, Prospective Studies Collaboration, 2009).

Like other chronic disorders, obesity has had an affinity towards individuals with particular demographics, socioeconomic status, and those negatively affected by the social determinants of health. The high prevalence of obesity in the U.S. has been associated with people of color, older individuals, and those residing within the southeastern states (CDC, 2012; Flegal, Carroll, Oglen, & Curtin, 2010; Taber, Stevens, Poole, Maciejewski, Evenson, et al., 2011; WHO, 2013). Income has also been an influential risk factor in determining obesity. A lower income has consistently corresponded to an increased occurrence of obesity, especially among people of color and women (CDC, 2012). To stabilize these trends in obesity and its discrimination towards particular groups, risk factors should be identified and acknowledged in order to create potential interventions and public policies that address obesity's multi-etiologic nature.

Significance of the Study

Obesity is a result of multiple risk factors that include diet and behavior. Its incidence has increased over the last four decades, leading to decreased quality of life and life expectancy (Jolliffe, 2011; Mokdad, Ford, Bowman, Dietz, Vinicor, et al., 2003; Peeters, Barendregt, Willekens, Mackenbach, Al Mamun, et al., 2003; Prospective Studies Collaboration, 2009; WHO, 2013). By analyzing risk factors, exposing obesity's interdependent risk factors may lead to the proposal of a more intimate and integrated policy and/or intervention.

LITERATURE REVIEW

Over the last two decades, there has been a drastic increase in the prevalence of obesity (CDC, 2011). Data from the CDC has shown that the dramatic rise in obesity began in the 1990s when there was less than 10 to 14 percent obesity. A decade later, there was a recorded 10 percent increase in the range of prevalence of obesity. In the most recent decade, 2010, obesity rose to a range of 20 percent to greater than 30 percent amongst adults, a ten to sixteen percent increase from two decades prior (CDC, 2011).

This two-decade increase has been linked to particular races, age groups, genders, locations, and behaviors (Flegal, Carroll, Ogden & Curtin, 2010). Researchers have found that non-Hispanic Black women were most often obese. With respect to age, these investigators have also observed that obesity was most prevalent amongst older men and middle-aged women, but in general, obesity was more evident and occurred at a faster rate in men (Flegal, Carroll, Ogden & Curtin, 2010). Although obesity has affected men at a faster rate, it has impacted a greater proportion of women (Flegal, Carroll, Ogden & Curtin, 2010).

Other demographics associated with the prevalence of obesity have included level of education, smoking status, and location.   Less educated smokers and Mississippi residents had higher instances of obesity than those without these characteristics. Not only does an individual's susceptibility tend to be more biased depending on their characteristics, the chronic, acute, and mental morbidities associated with obesity in the U.S. tend to be repetitive amongst this same group (Mokdad, et al., 2003).

The Morbidity and Mortality Associated with Obesity

Obesity and its corresponding morbidity and mortality have had an overall increased occurrence across America. Increases in these facets of obesity within adults have also been experienced among the youth, according to a six-year nationwide study. Amongst children, this increase was proposed to be a result of television watching, milk and fruit juice consumption, participation in sports, race, and socioeconomic status (Taber, et al., 2011).   The presence of obesity and its associated morbidity and mortality have also had an affinity towards particular states and cities. Similar to the evaluation of different countries, the direction and magnitude of the prevalence of obesity are dependent on the individual's residency, regardless of age range (Taber, et al., 2011). Researchers have discovered an exceedingly high percentage of obese individuals in the southeast United States. With increased incidence due to age, gender, residence, and the resultant effects of its morbidity and mortality, legislative attempts to moderate this rise have failed.

Risk Factors: The Effects of Unemployment and Poverty

Two fiscally associated obesity risk factors include unemployment and poverty rates. The Bureau of Labor Statistics has defined unemployment as, "individuals who are unemployed, but are capable of working" (United States Department of Labor, 2009). Poverty is based on family size and income to determine the financial threshold. When the individual/value is below the qualifying threshold, he or she is defined as impoverished (United States Census Bureau, 2011). According to a previous study, one of the effects of compromised financial circumstances caused by unemployment and poverty on obesity has been seen in terms of poor food selection (Weaver, Hayes, Unwin & Murtagh, 2008).

Increased costs in food serve as an additional financial constraint amongst those with a limited budget. Less expensive and less nutritive options are more tangible to the limited budget, which justifies the unemployed and impoverished intake of higher quantities of these foods that predispose them to obesity (Weaver et al., 2008). Although unemployment confines individuals to cheaper food options, Weaver et al. (2008), proposed that employment also exaggerates increases in obesity. By working long hours, employed individuals have easier access to worse diets because of time restraints. In addition, Weaver et al. (2008), suggested that a more sedentary job promotes a less active lifestyle, usually resulting in less fit employees.

The trend concerning employment only holds true to a certain extent. Decreases in BMI have also been experienced among those with an income greater than or equal to 130 percent above the poverty line (Joliffe, 2011). Assuming that being overweight predisposes an individual to obesity, it is noteworthy to state that the poor have historically had a higher prevalence and severity of being overweight (Jolliffe, 2011). Guided by income, both unemployment and poverty rates have had a negative association with the healthy eating capacity of non-Hispanic Blacks compared to non-Hispanic Whites (Wang & Chen, 2011). Non-Hispanic Blacks are almost 50 percent less likely to have a high healthy eating index. Wang and Chen (2011), associated these discrepancies with individuals being either impoverished and/or unemployed, which could predispose them to obesity.

Risk Factor: The Effects of Educational Attainment

A study evaluating food choices concluded that those who are more educated usually make wiser food choices (Wang & Chen, 2011). Opting out of these choices has resulted in premature mortality among less educated individuals, especially in recent years (Cutler, Lange, Meara, Richards-Shubik, & Ruhm, 2011). The occurrence of obesity increased from one to six percent for both men and women with less than a high school education. Additionally, over the last two decades, the risk of obesity doubled when educational level was less than a high school education (Cutler, Lange, Meara, Richards-Shubik, & Ruhm, 2011).

Risk Factors: The Effects of Alcohol Consumption and Smoking Status

Drinking and smoking have consistently been associated with weight. It has been assumed that there is an increase in obesity that correlated to both habits. In certain studies, the use of alcohol and drugs actually lowered the risk of obesity compared to normal weight individuals (Gearhardt, Harrison & McKee, 2012). With respect to alcohol, alcohol consumption has been commonly thought to increase the risk of obesity. Chronic consumption allows for increased caloric intake, abnormal fat metabolism, alcohol abuse, and poor eating habits (Davis, 2010). In addition, BMI tended to be higher in current drinkers when they consumed greater than or equal to four drinks at a low frequency and low volume (Breslow & Smothers, 2009).

Regarding smoking, smoking cessation has been more than likely to result in additional weight gain than smoking itself (U.S. National Library of Medicine, 2013). Studies have suggested that nicotine, the key chemical agent, can reduce weight gain because of its anti-inflammatory capabilities that may suppress the inflammation associated with obesity, and yield a decrease in weight (Lakhan & Kirchgessner, 2011). In addition to suppressing pro-inflammatory pathways, nicotine has been believed to be an appetite suppressant (Lakhan & Kirchgessner, 2011).

Risk Factors: The Effects of Participation in Physical Exercise with respect to Violence/Crime

Of the risk factors analyzed in this study, participation in physical exercise has been the only factor to have a negative association with BMI (Sallis & Glanz, 2009). Logically, if exercising results in a decline in calories, it is understandable that the absence of this risk factor would indirectly correspond to an increased incidence of obesity. Participation in physical activity usually occurs after a personal, negative experience by a close friend or relative with the complications of weight, in addition to concerns regarding physical appearance. Even with the additional motivation, increases in weight, age, and problems due to arthrosclerosis hinder rigorous participation in physical exercise, resulting in excess body fat, even amongst different races (Weaver, Hayes, Unwin, & Murtagh, 2008). The effect of race becomes exaggerated by lack of participation in physical exercise. Non-Hispanic Blacks are approximately 30 percent less likely to participate in physical exercise, regardless of their nutritional knowledge and awareness, self-efficacy concerning a healthy diet, food choice, income, and education (Wang & Chen, 2011). This finding was possibly a result of the unfavorable environment(s) where most minorities typically reside (Sallis & Glanz, 2009).

We are convinced that people may not participate in physical exercise frequently partly because of their environment, like the presence of violence/crime and changes in scenery. With changes to the landscape like closer proximity to recreational facilities and such destinations of interests as shopping centers and schools, physical activity have the potential to increase. The current remoteness of places of

interest has promoted vehicular transportation, which has reduced participation in physical activity (Sallis & Glanz, 2009). Although there are infrequent publications on the effect of violence/crime and the participation in physical activity, it is speculated that increased presence of violence/crime hinders the motivation to participate in physical activity.

Risk Factor: The Effects of the Average Allotment of Food Stamps per Month

Non-Hispanic Blacks were less likely to consume a more expensive, high quality meal because they were approximately 50 percent less likely to make healthy food choices, and had lower intentions of making improvements to their diet. This action could be a result of wanting to be more efficient with food assistance resources, i.e., increasing the quantity of groceries for less fiscal resources. According to Sallis and Glanz (2009), this action is also due to their nutritional environment, which is saturated with fast-food markets and supermarkets lacking a variety of healthful choices, limiting their self-efficacy in selecting a more appropriate diet. Due to their incognizance about food choices, possibly guided by lack of income, the BMI amongst this group has been higher, in comparison to higher- income persons (Wang & Chen, 2011).

Risk Factors: The Effects of Physician Care and Hospital Care Costs

Compared to normal weight and overweight individuals, obese clients visit the physician more often because of acute complications from these chronic conditions or pre-/post-operative surgery (Pearson, Bhat-Schelbert, Ford & Mokdad, 2009). These conditions gave rise to spending more time with the physician, as well as being prescribed more medications per visit, resulting in higher physician associated costs.

Increased costs of hospital care have been presumed to be a result of increased severity in disease by the time services are rendered. Through the analysis of a free clinic, it was concluded that clients, irrespective of BMI, most often utilize these services for a chronic medical condition. Out of those that solicited these services, approximately one-third were obese (Notaro, Khan, Bryan, Kim, Osunero, et al., 2011).

Theoretical Framework- Social Cognitive Theory

The social cognitive theory is based on the term reciprocal determinism to produce a particular behavior, in which environmental, behavioral, and personal factors all influence one another (Glanz, Rimer & Viswanath, 2008). The interdependence of the three factors relies on the following five concepts of the social cognitive theory: psychological determinants, observational learning, environmental determinants of behavior, self-regulation, and moral disengagement (Glanz, Rimer & Viswanath, 2008). This concept was used to explain the relationship between obesity and possible risk factors encompassed within the personal, environmental, and behavioral realm. By being considerate of the different dynamics of obesity in relation to the social cognitive theory, a more comprehensive and plausible solution to regulate its increases can be suggested after the significant and important risk factors are determined.

METHODOLOGY

Overview/ Primary Aim

The primary gpal of this study was to determine if the prevalence of obesity in the United States for years 2000 to 2006 was affected by: unemployment rate, poverty rate, the percentage of the population that either had less than nine years of an education, entered high school and did not receive their high school diploma, or entered high school and received their high school diploma, alcohol consumption per capita,

percent of smokers in state, participation in physical exercise, violence/ crime rate, average allotment of food stamps per month, cost of hospital or physician care, respectively, and calendar year. Data was collected from different federal surveys that were inclusive of the risk factors for this study. To achieve this study's objective, Partial-Least Square (PLS) regression and Artificial Neural Network (ANN) models were constructed to determine the statistical significance and relative importance of the thirteen risk factors.

PLS regression has the ability to model linear relationships between sets of independent and dependent variables (Barker & Rayens, 2003). In this scenario, the independent variables are analogous to the risk factors and are used to make predictor measures for their relationship with the obesity outcome (Boulesteix & Strimmer, 2007). This method was chosen over multiple linear regression because it tolerates highly correlated relationships among independent variables, the so-called collinearity, even among a small sample size (Zou, Yun, & Zhang, 2009).

ANN is a multifaceted and analytical approach which uses input, hidden, and output layers to analyze nonlinear rank of importance among input and output variables. Data for the independent variables (or the risk factors) proceeds from the input layer to be analyzed in the sequential layers, where activation function(s) and weights are taken into account. Its error computations, via back propagation, ensure optimal results because of the minimal residuals between the observed and the expected results (Chen, 2010). Both models were used for this study in a collaborative, linear and nonlinear approach to analyze the statistical significance and rank of importance of these risk factors.

Research Design/Data Collection

In the analysis of significance and relative importance of obesity's risk factors for years 2000 to 2006 in the U.S., obesity was the dependent variable. The null hypothesis was that there is no association between obesity rate (the number of individuals with a BMI greater than or equal to 30 kg/m^2 per 1,000 individuals) and the aforementioned risk factors. This multi-methodological study used statewide census survey data to analyze the significance and importance of obesity's possible risk factors for years 2000 to 2006. Due to the data collection methods and use of results, this study was considered cross-sectional and correlational.

State-level data from years 2000 to 2006, acknowledging the previously-mentioned risk factors and obesity rate, was gathered, resulting in sample size (n) of 350 (seven years x 50 states). To unveil all significant relationships between obesity rate and its potential risk factors (See Table 1), this retrospective study used data that was consolidated from surveys by: the United States Department of Labor, Centers for Disease Control and Prevention (CDC), United States Department of Agriculture- Food and Nutrition Services (USDA), United States Census Bureau, National Institute on Alcohol Abuse and Alcoholism, Federal Bureau of Investigation (FBI), and the Center for Medicare/Medicaid Services (CMS). Data were consolidated into a continuous file by variable, state, and calendar year and exported to the statistical package, XLSTAT, for the PLS regression model as well as to the IBM Statistical Package for the Social Sciences (SPSS) for the ANN model.

Table 1.

Variable Names and Criteria for the Risk Factors Potentially Associated with Obesity Rate

Variable Abbreviation	Variable Description	Measurement Scale	Source
Unemploy	Unemployment Rate: a measure of prevalence of unemployment (the amount of people that are without jobs, but are actively seeking employment within the last few week)	Measured by dividing the number of unemployed individuals by all individuals currently in the labor force	United States Census Bureau
PovRate	Poverty Rate: a percentage of people in poverty per calendar year by the Current Population Survey. If the family income before taxes and without capital gain/noncash benefits is less than threshold, the entire family is impoverished	Measured annual poverty rates based on the summation of reported annual income reported a few times a year divided by the sum of poverty thresholds at monthly intervals, resultant of changes in family composition	Office of Management and Budget's Statistical Policy Directive 14.
Less9ED	Percent of Population Less than nine Years of Education: this measure of educational attainment referred to the highest level of education that an individual has completed	The American Community Survey (ACS) is annually completed by about 3.5 million randomly selected households in the US. The question on the survey reads: "What is the highest degree or level of school this person has completed? "	United States Census Bureau
HSch_Nod	Percent of Population Who Entered High School But Did Not Receive Diploma	For description, please see Less9ED	United States Census Bureau
HSch_Yes	Percent of Population Who Received High School Diploma	For description, please refer to Less9ED	United States Census Bureau
Alcohol	Alcohol Consumption per 1,000 Gallons: The liters of pure alcohol per capita by state were computed as the sum of alcohol production and imports, divided by the adult population (aged 15 years and older)	Calculated from official statistics on production, sales, import, and export, taking into account stocks whenever possible.	National Institute of Alcohol Abuse and Alcoholism

Cigarette	Percentage of Smokers in State: was measured by conducting telephone interviews by the Behavioral Risk Factor Surveillance System (BRFSS)	More than 350,000 adults are interviewed each year and asked the question, "Are you a current smoker?"	CDC
P_Exerci	Percentage who said they Participated in a Physical Activity per Month	More than 350,000 adults are interviewed each year by telephone using BRFSS and asked the question, "During the past month, did you participate in any physical activities?"	CDC
V_Crime	Violence Crime Rate per 100,000 People: measured the number of incidents such as murder, rape, and aggravated assault that occurred annually.	Measured by the crime rate per 100,000. The calculation of crime rates uses data that is obtained either from criminal justice systems or from public survey	FBI
FoodStamp	Average Food Stamp Allotment Per Month	Averages derived from total benefits divided by total participation	USDA
HospCare	Cost of Hospital Care/ Healthcare Expenditures	The value of hospital care is measured by total net revenue, which equals gross patient revenues (charges), less contractual adjustments, bad debts, and charity care	CMS
PhysCare	Cost of Physician Care/ Healthcare Expenditures	Measured by total net revenue, which equals gross patient revenues (charges), less contractual adjustments, bad debts, and charity care	CMS
YRCode	Year: for the seven year study, each year was representative of its own increment in order to determine if year was a risk factor for obesity rate	Measured by calendar year	Not Applicable

Data Processing and Analysis

Data from the federal surveys were organized in Microsoft Excel alphabetically by state and in chronological order before being exported to be analyzed through PLS (XLSTAT, 2012) and ANN (SPSS, Version 20). In this study, the significance of the 13 risk factors was examined, in relation to obesity rate at the 0.05 significance level via PLS.

The rankings of the risk factors were also assessed based on the relative importance via ANN. The connectivity of the input, hidden, and output layers revealed the correlation between the risk factor in relation to the output variable, obesity rate. It does so by the use of functions like hyperbolic, sigmoid, or identity functions and pre-calculated weights, which define the tri-layered network (SPSS, 2008). Normalized importance was calculated by creating a ratio of importance of the independent variables independently to smoking status, the highest value of normalized importance (Chen, 2010). This ratio generated a percentage that produced the normalized importance of each variable. Since PLS is also capable of ranking independent variables through the variable importance of projection (VIP), it was used to evaluate the consistency of rankings.

In this study, the hyperbolic tangent and identity activation functions were used in a feed-forward, multilayer perceptron method between the input to hidden and hidden to output layers, respectively. In addition to the technique of feed-forward, minimal error was achieved by using the back propagation algorithm. The use of two different procedures along with input, hidden, and output layers characterized the model as multilayer perceptron (Chen, 2010). By using the ANN model, data can transition from input to output to rank the order of relative importance with high accuracy, according to the criteria of mean absolute percentage error (MAPE) and root mean square error (RMSE).

STUDY RESULTS

Summary Statistics and PLS Goodness-of-Fit

Analysis of 13 risk factors over seven years within all 50 states resulted in a sample size of 350 data points. PLS regression analysis resulted in a goodness-of-fit statistic, R-squared value, of 0.540. Based on this value, more than half (54 percent) of the variation in obesity rate is explicable by the risk factors. The mean square error (MSE), which reveals the variability between the observed and predicted value, was 4.713. This miniscule value also correlated to a minimal root mean square error (RMSE) (RMSE = 2.17). Both are also indicative of good model fitting and accurate result. Additionally, there was 89 percent accuracy of the trained data and approximately 77 percent accuracy with tested data.

Statistical Significance According to PLS

PLS regression is capable of detecting statistical significance. Insignificant variables had a confidence interval of standardized coefficients that crossed the x-axis (zero value) and were both negative and positive in confidence limits. In contrast, significant variables were either positive or negative. The following risk factors positively influenced the obesity rate: unemployment rate, poverty rate, percentage of smokers, average allotment of food stamps per month, and the cost of hospital and physician care. The risk factor negatively associated with obesity rate was alcohol consumption per capita. Interpretively, the relationship between increasing obesity rate correlates to a decrease in alcohol consumption per capita; all other significant risk factors increase with obesity rate, where an increase in the respective factor corresponds to an increase in obesity rate.

Figure 1.

Ninety-five (95) Percent Confidence Intervals of the Standardized Coefficients for Risk Factors Associated with Obesity Rate
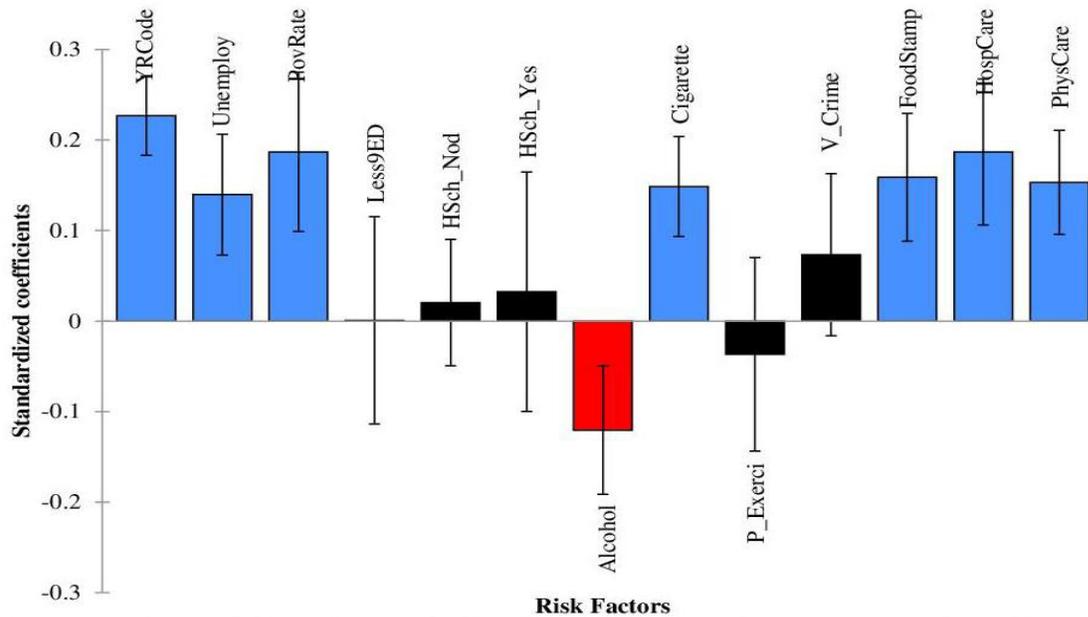
Seven risk factors (shown in blue) shared a positive association with obesity while alcohol consumption per capita (shown in red) is the only significant, negatively related risk factor. All others risk factors were insignificant (shown in black).

Relative Importance of Risk Factors According to ANN

This study used ANN because of its capability to rank risk factors in order of normalized importance. The normalized importance of the risk factors of interest is a hierarchal ranking method that compares the ratios of relative importance among risk factors (Chen, 2010). The greatest value of relative importance (percentage of smokers in state) from the individual risk factors is used to calculate normalized importance as a reference value (denominator). The risk factors most relevant to obesity had minimal difference between comparative values of the normalized importance ratio.

Of the potential risk factors relevant to obesity, more than 11 had greater than 50 percent relative importance to obesity rate. These risk factors include: unemployment rate, poverty rate, all levels of educational attainment, alcohol consumption per capita, percentage of smokers in state, average allotment of food stamps per month, cost of physician and hospital care, respectively, and calendar year. These results did not reveal a thematic trend for the risk factors evaluated.

Figure 2.

Graphical Representation of the Normalized Importance of the Risk Factors Associated with Obesity Rate
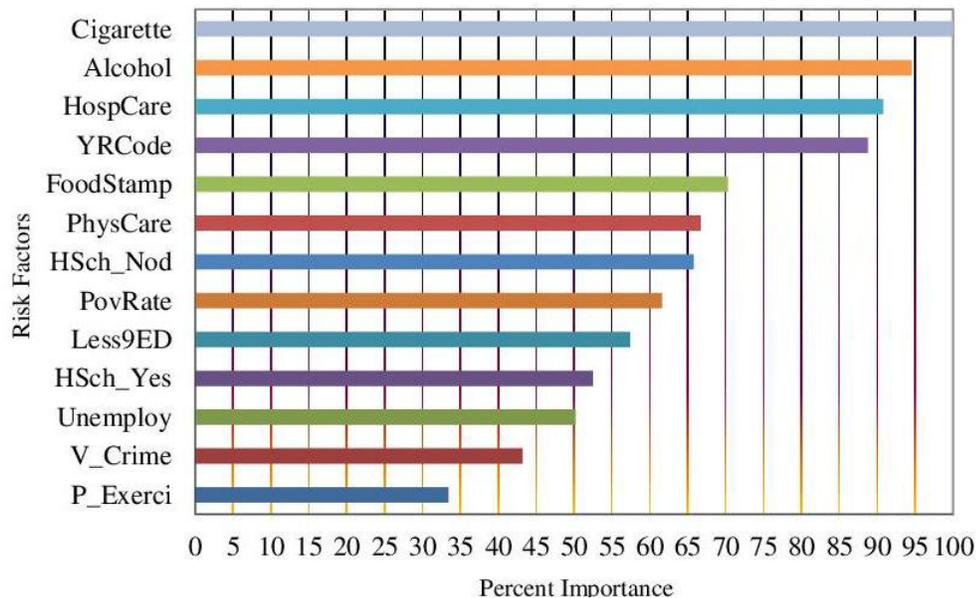
DISCUSSION and CONCLUSION

Significant risk factors associated with obesity over the seven-year period were personal (average allotment of finances for food stamps, hospital care, and physician care), behavioral (fraction of smokers in state and alcohol consumption per capita), and environmental (poverty rate, unemployment rate, and calendar year). Each significant variable had at least 50 percent of relative importance with respect to obesity. Most of the risk factors significantly associated with obesity were positively correlated, meaning that a percentage decrease of smokers by state would result in a decrease in obesity. Other risk factors following this trend were: unemployment rate, poverty rate, average allotment for food stamps per month, and the cost of hospital care and physician care. The only negatively-associated risk factor was alcohol consumption per capita, which had the second highest rank of normalized importance. Irrespective of its importance, the negative association definitively meant that an increase in alcohol consumption per capita yielded a decrease in obesity and vice versa.

Obesity has been rising consistently with time (CDC, 2010; Flegal, Carroll, Ogden & Curtin, 2010; Jolliffe, 2011). This increase has been explicable due to the decreased proximity of places of interest and technological advancements. Technological advancements at work and home promote a sedentary lifestyle because of television watching and computer use (Taber, et al., 2011; Weaver, Hayes, Unwin & Mutagh, 2008). In addition, the distance of work, school, and shopping centers from residences promotes vehicular transportation, an additional technological advancement, which increases the amount of time in a sedentary stance (Sallis & Glanz, 2009).

Unemployment, poverty, and allotment of food stamps per month reduce the accessibility of healthy food options because of financial restrictions. Decreased income has been correlated to increases in BMI (Jolliffe, 2011). Limited fiscal resources limit the healthy eating capacity of the individual (Wang & Chen,

2011). In addition, the limitations of finances prompt individuals to purchase higher quantities of less nutritive foods because they are more feasible, accessible, and the recipients are usually incognizant of what qualifies as nutritive (Sallis & Glanz, 2009; Weaver, Hayes, Unwin, & Mutagh, 2008).

Smoking status and alcohol consumption were two of the more historically reviewed risk factors that yielded inconclusive results. This study determined that there is a positive association with smoking status and negative association with alcohol consumption. Smoking corresponded to a lower BMI because of the metabolic action of nicotine (Lakhan and Kirchgessner, 2011). An increase in BMI, resultant of chronic smoking, followed by smoking cessation, has also been found to promote obesity (U.S. National Library of Medicine, 2013). Increases in BMI were also explained by increased food intake with increased participation in smoking (Wehby, Muray, Wilcox & Lie, 2012). The negative association between alcohol consumption and BMI is due to the relationship of food with concentration of alcohol (Vanden, Heuvel & Dineen, 2012). They proposed that blood alcohol concentration (BAC) is inversely associated with food consumption, which meant that a lower BAC correlated to higher food consumption.

The final two risk factors associated with obesity were related to hospital and physician care expenses. The positive association between both expenses and BMI are thought to be due to more frequent visits to physicians and hospitals. Increased hospital costs were due to heightened surgical requirements and disease severity due to coexisting chronic conditions (Notaro, et al., 2011; Pearson, Bhat-Schelbert, Ford & Mokdad, 2009). Increased BMI increased the number of prescribed medications and cost because of the additional complications from obesity and its comorbid disorders (Pearson, Bhat-Schelbert, Ford & Mokdad, 2009. In conjunction with one another and in accordance with previous studies, the significant and important risk factors of this study have promoted obesity and should be justifiably targeted for future policies and interventions.

Within the last few decades, obesity has increased amongst children and adults across the United States (CDC, 2011). Being obese as a child can create additional complications as an adult, in addition to the chronic diseases that come with age (Taber, et al., 2011). Changes in technology, accessibility of food, and relative proximity of places of interest promote a less active lifestyle with consumption of less nutritive foods, and promote obesity. The prevalence of obesity and its associated morbidity and mortality have been biased towards particular regions, individuals with specific demographics, and levels of socioeconomic status (CDC, 2011; Flegal, Carroll, Ogden & Curtin, 2010; WHO, 2013).

In agreement with this study, obesity is linked to more than one risk factor and therefore, policies and interventions should be created to address more than one of these factors. To moderate inclines in obesity, researchers must acknowledge obesity's multi- etiologic nature and be cognizant of the role of the social cognitive theory. By receiving conclusive data from retrospective analysis, other longitudinal study designs can be used to validate study findings on a closer scale.

Risk factors having an insignificant relationship with respect to obesity include educational attainment, participation in physical exercise, and the presence of violence/crime. Previous studies have suggested a relationship between these risk factors and high BMI (Prospective Studies Collaboration, 2009; Pan & Yeh, 2008; Dixon, Dixon, & O'Brien, 2003; Flegal, Graubard, Williamson, & Gail, 2007; Peeters, et al., 2003). Unlike previous studies, this study negated any significant relationship due to educational

attainment, specifically, less than nine years or possession/lack of high school diploma, the presence of violence/crime, or the amount of physical activity.

Study Strengths and Limitations

The major strength of this study is related to the models of analysis.   The two models, PLS and ANN, yielded highly accurate results because they validate one another. Analysis of significance of risk factors can be accomplished by multiple linear regression that allows researchers to evaluate multiple independent variables in relation to the dependent variable, but requires that the independent variables be uncorrelated among themselves. The benefit of PLS over multiple linear regression is that PLS can tolerate collinearty. This study required the suitable PLS regression model to handle the collinearty problem by using the principal component technique to regroup highly correlated variables into a few components. An additional benefit of PLS is that it ranks the relative importance of risk factors through variable importance in the projection (VIP). With respect to ANN, by comparison, these two processes can validate one another. Finally, since ANN uses a feed-forward back propagation method by testing and training sample data, the model is thought to be highly accurate.

A limitation of the study included convenience sampling. Since this study used secondary data, it is possible that those that responded to each federal survey were part of a convenience sample. Through convenience sampling, it is possible that respondents are not truly representative of the prevalence of the neither risk factors nor obesity for the seven year time period of interest. Under this notion, there may be limited external validity of study results because of possible selection bias. In addition, the survey respondents completing the federal surveys were not necessarily the same group each time. Therefore, the inclusion and exclusion criteria for the surveys may have varied and skewed results. To eliminate the study limitations caused by choice of study design, a prospective study would have been suggested. This procedure would have allowed declarations on temporality, but would have not corrected for causality. Thus, the limitations of a prospective study outweigh those of a cross-sectional study because they include loss to follow up, financial burdens, as well as require a large investment of time.

Public Health Implications

Irrespective of age, gender, location, socioeconomic status, and other factors, the occurrence of obesity has been steadily rising across the United States (Taber, Stevens, Poole, Maciejewski, Evenson, et al., 2011). Federal, noninvasive treatments of obesity propose that increased physical exercise and changes in diet may entice other behavioral changes to decrease the percentage of body fat (U.S. National Library of Medicine, 2013). In addition, while studies have proposed increased physical activity results in decreased percentages of body fat, this study has proven that, with respect to the other possible risk factors, physical activity is not significant in regulating obesity.   It is probable that these mandates and suggestions have been ineffective at regulating these trends of mandate limitations, because they only addressed the accessibility of healthy and non- nutritive foods, targeted areas with high violence/crime rates, and individuals with a low education level, and have suggested that physical activity could serve as a remedy.

Studies have shown that an association exists between poverty, income, and unemployment amongst other risk factors. This study agreed with previous results, to an extent, and built upon previous findings in terms of the significance of risk factors of interest. As proven by this analysis, obesity is influenced by personal, behavioral, and environmental factors. Knowing the existence of these associations may be extremely pertinent in developing future interventions and public policies. The risk factors possibly

explaining the highest amount of variation in the risk factor-outcome relationship were: calendar year, poverty rate, costs of hospital care, average allotment of food stamps per month, costs of physician care, proportion of cigarette smokers in state, and unemployment rate. Alcohol consumption per capita explains very little of the variation with obesity rate through an inverse relationship. Results tend to agree with what was expected because those who are at the highest risk of suffering from obesity include: women, those of an ethnic background, lower income individuals, and older populations (Flegal, Carroll, Ogden & Curtin, 2010). These individuals have an affinity to be characterized by the significant risk factors of this study. From study results and in accordance with most studies referenced, the risk factors that should be targeted for public policies and interventions are calendar year, poverty rate, costs of hospital and physician care, alcohol consumption per capita, average allotment of food stamps per month, and percent of smokers.

REFERENCES

Barker, M. & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, *17*(3), 166-73.

Boulesteix, A.L. & Strimmer, K. (2007). Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. *Briefings in Bioinformatics*, *8*(1), 32-44.

Centers for Disease Control and Prevention. (2011, September). Healthy weight- it's not a diet, it's a lifestyle. Retrieved from http://www.cdc.gov/healthyweight/assessing/ bmi/adult_bmi/index.html

Centers for Disease Control and Prevention. (2012, August). Obesity and overweight for professionals*: Data and statistics*. Retrieved from http://www.cdc.gov/obesity/ data/adult.html

Centers for Disease Control and Prevention. (2012, May). Public Health Genomics: Genomics and Health. Retrieved from http://www.cdc.gov/genomics/resources/diseases/ obesity/index.htm

Chen, C.K. (2010). Curriculum assessment using artificial neural network and support vector machine modeling approaches: a case study. *IR Applications*, *9*, 1-24.

Cutler, D.M., Lange, F., Meara, E., Richards-Shubik, S., & Ruhm, C.J. (2011). Rising educational gradients in mortality: the role of behavioral risk factors. *Journal of Health Economics*, *30*, 1174-87.

Davis, S. (2010). Facts on alcohol and obesity. Retrieved from http://www.livestrong.com/article/69121-alcohol-obesity/

Dixon, J.B., Dixon, M.E., & O'Brien, P.E. (2003). Depression in association with severe obesity- changes with weight loss. *Archive of Internal Medicine*, *163*, 2058-65.

Flegal, K.M., Carroll, M.D., Ogden, C.L., & Curtin, L.R. (2010). Prevalence and trends in obesity among US adults, 1999- 2008. *Journal of the American Medical Association*, *303*(3), 235-41.

Flegal, K.M., Graubard, B.I., Williamson, D.F., & Gail, M.H. (2007). Cause specific excess deaths associated with underweight, overweight, and obesity. *Journal of the American Medical Association*, *298*(17), 2028-37.

Gearhardt, A.N., Harrison, E.L.R., & McKee, S.A.(2012). Does co- morbid depression alter the inverse relationship between obesity and substance use disorders? *Drug and Alcohol Dependence*, *124*, 1-4.doi: 10.1016/jdrugalcdep.2012.01.002.

Glanz, K., Rimer, B.K., & Viswanath, K. (2008). *Health behavior and health education: Theory, research, and practice* (4th ed.), (pp. 169- 88). San Francisco, CA: Jossey- Bass.

Jolliffe, D. (2011). Overweight and poor? On the relationship between income and the body mass index. *Economics and Human Biology*, *9*, 342-55.

Krukowski, R.A., Smith-West, D., Harvey- Berino, J., & Prewitt, E. (2010). Neighborhood impact on healthy food availability and pricing in food stores. *Journal of Community Health*, *35*, 315-20.

Lakhan, S.E. & Kirchgessner, A. (2011). Anti- inflammatory effects of nicotine in obesity and ulcerative colitis. *Journal of Translational Medicine*, *9*(129), 1-10. doi:10.1186/1479-5876-9-129.

Medical Encyclopedia, A. D. A. M. (2010, July). Body mass index. Retrieved from http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0004463/

Mokdad, A.H., Ford, E.S., Bowman, B.A., Dietz, W.H., Vinicor, et al (2003). Prevalence of obesity, diabetes, and obesity- related health factors, 2001. *Journal of the American Medical Association*, *289*(1), 76-9.

Notaro, S.J., Khan, M., Bryan, N., Kim, C., Osunero, T., et al (2011). Analysis of the demographic characteristics and medical conditions of the uninsured utilizing a free clinic. *Journal of Community Health*, *37*(2), 501-6. doi: 10.1007/s10900-011-9470-7.

Pan, W.H. & Yeh, W.H. (2008). How to define obesity? Evidence- based multiple action points for public health awareness, screening, and treatment: an extension of Asian- Pacific recommendations. *Asian Pacific Journal of Clinical Nutrition*, *13*(3), 370-4.

Pearson, W.S., Bhat- Schelbert, K., Ford, E.S., & Mokdad, A.H. (2009). The impact of obesity on time spent with the provider and number of medications managed during office- based physician visits during a cross-sectional, national health survey. *BMC Public Health*, *9*(436), 1-7.

Peeters, A., Barendregt, J.J., Willekens, F., Mackenbach, J.P., Al Mamun, A., et al (2003). Obesity in adulthood and its consequences for life expectancy: A life- table analysis. *Annals of Internal Medicine*, *138*(1), 24-33.

Prospective Studies Collaboration. (2009). Body- mass index and cause- specific mortality in 900,000 adults: collaborative analyses of 57 prospective studies. *The Lancelet*, *373*, 1083-96.

Sallis, J.F. & Glanz, K. (2009). Physical activity and food environments: solutions to the obesity epidemic. *The Milbank Quarterly*, *87*(1), 123-54.

Taber, D.R., Stevens, J., Poole, C., Maciejewski, M.R., Evenson, K.R., et al (2011). State disparities in time trends of adolescent body mass index percentile and weight- related behaviors in the United States. *Journal of Community Health*, *32*(2), 335-43. doi:10.1007/s10900-011-9442-y

Tsuguane, S. (2012). Alcohol, smoking, and obesity epidemiology in Japan. *Journal of Gastroenterology and Hepatology*, 121-6. doi: 10.1111/j.1440-1746.2011.07021.x

United States Census Bureau. (2011). Retrieved from http://www.census.gov/ hhes/www/poverty/methods/definitions.html

United States Department of Labor. (2009). Retrieved from http://www.bls.gov /cps/cps_htgm.htm

U.S. National Library of Medicine. (2013). National Center for Biotechnology Information: Obesity. Retrieved from http://www.ncbi.nlm.nih.gov/pubmedhealth/

Wang, Y. & Chen, X. (2011). How much of racial/ethnic disparities in dietary intakes, exercise, and weight status can be explained by nutrition- and health- related psychosocial factors and socioeconomic status among US adults. *Journal of the American Dietetic Association*, *111*, 1904-11.

Weaver, N.F., Hayes, L., Unwin, N.C., & Murtagh, M.J. (2008). "'Obesity' and 'Clinical Obesity'": Men's understandings of obesity and its relations to the risk of diabetes: A qualitative study. *BioMedCentral Public Health*, *8*(311), 1-8.

Wehby, G.L., Murray, J.C., Wilcox, A. & Lie, R, T. (2012). Smoking and body weight: evidence using genetic instruments. *Economics and Human Biology*, 113-26. doi: 10.1016/j.ehb.2011.09.002.

World Health Organization. (2013, March). *Obesity and overweight*. Retrieved from http://www.who.int/mediacentre/factsheets/fs311/en/index.html

Zou, Z., Yun, Y., & Zhang, W. (2009). Study on relationship between Beijing's living water demand and urbanization with simple partial least-squares regression. *Services and Sciences, Management and Engineering*, 526-30. doi: 10.1109/SSME.2009.140