Vol. 10(8), pp. 281-286, August 2018
DOI: 10.5897/JPHE2017.0946
Article Number: 3E0052857810
ISSN: 2141-2316

http://www.academicjournals.org/JPHE

ACADEMIC JOURNALS
expand your knowledge

Journal of Public Health and Epidemiology

Full Length Research Paper

Statistical model of weight outcome among college students

Maxwell Seyram Kumah[1*], Awuitor Godwin Kofi[2], Dorkenoo Bridget Aku[2] and Kumi John Newton[1*]

[1]Mathematics and ICT Department, St. Teresa's College of Education, Box 129, Hohoe, Ghana.
[2]Mathematics and ICT Department, Akatsi College of Education

Received 6 June, 2017; Accepted 31 July, 2017

The issue of weight outcome is a common problem across all ages in the world, regardless of being a male or female. Studies show that its cause is more associated with insufficient physical activities due to technological addicts and improper eating habits. However, this study sought to find out whether the predictors by authorities are in existent among college students in this part of the country. In a cross sectional survey which was a cohort study on 308 subjects, predictors like gender, age, eating habits and physical exercise were tested on weight outcome as the dependent variable (DV) that was categorized into two (normal and abnormal weight), which is *dichotomous*. The abnormal weight is made up of underweight, overweight and obese. Test of association was computed and Pearson Chi-square value was found to be $x^2 = 6.662$ with significance value of 0.01 < 0.05. The result revealed that since the significance value is less than the alpha value, then gender is associated with a weight problem. The logistic regression model (LRM) was used to analyze the epidemiological behaviour of the problem. The Hosmer and Lemeshow test computed, displayed a significance value of 0.291 which is by far greater than the 5% significance level and hence affirm that the data used was good to fit the model. It was also maintained that overall, 63.0% of the subjects were correctly classified. It was further revealed from the LR table that gender had p – value 0.014: C.I = 1.288 – 9.292; age, p = 0.001: C.I. = 2.173 – 8.132; eating habits, p = 0.021: C.I = 0.323 – 0.913 and physical exercise, p = 0.001: C.I. = 2.303 – 20.232, this establishes that each of these independent variables IVs is <0.05, hence significant to predict weight outcomes (problem).

Key words: Weight outcome, physical exercise, eating habits, odds ratio, risk.

INTRODUCTION

The prevalence of weight problem is now a common issue across all age groups in the world. According to Laquatra (2004), the cause of this increase prevalence is in two folds: (1) food is more readily available to everyone and (2) the physical activities of people are reduced with increasing urbanization and economic development leading to high energy gain from dietary intake. Observations made by the research further explain why many people in Ghana today pay much attention to work/studies and forget about exercise and in turn take

*Corresponding author. E-mail: Stiminca@yahoo.com.

more fast foods which are full of fat and oil. Unfortunately, these foods are mostly consumed in the night, which obviously is divergent to good health practices. As earlier stated, weight problem is alarming all over the world. According to Popkin (1994), an Africa country like South Africa who is experiencing economic transition shows an increase in overweight and obese, this is due to their economic condition changes. WHO (2000) also reported that about 6 to 8% of the population in Nigeria is obese and over 6% men and 13% women in Brazil are obese. Studies have also shown that apart from being overweight or obese, underweight problem is as equal as overweight and obese problems. Its cause was found to be associated with malnutrition among others and this problem is as a result of weight loss (Swinburn et al., 2004). The results of such weight-loss activities might lead to the development of under-nutrition and underweight (Keski-Rahkonen, 2003). Malnutrition is a condition caused by improper diet or nutrition, and can occur in conjunction with both under and over consumption of calories (World Food Programme, 2008; Rahman and Hakim 2016). It was also reported by Choi (2013) that negative attitudes towards obesity and socio-cultural preferences for thinness can even tempt persons who are already underweight to attempt weight control and both under-nutrition and overweight occur in institutional care. Body mass index (BMI) categorizes individuals as underweight (<18.5 kg/m^2), normal weight (18.5 to 24.9 kg/m^2), overweight (25.0 to 29.9 kg/m^2) and obese (≥30.0 kg/m^2) (Laquatra, 2004; WHO, 2013; Rahman and Hakim 2017), showing that both underweight and overweight/obesity represent serious public health challenges worldwide. Research has shown that Ghana is one of the African countries that made considerable progress over the past few decades'advancements in improving the health and well being of its citizens (Khawaja, 2008). In spite of this progress, there is still a need to educate more individual on problems especially, the young ones. Young ones because, observation of the lifestyle of people within the ages of 18 – 25 years, showa that they have bad eating habit and find it difficult to take part in physical activities. For example, they are interested in picking taxi over walking distance. It is against this background that we want to investigate the possible variables that are associated with weight problems.The scope of this study was delimited only to college students in the Hohoe Municipality.

METHODS

Study area

Hohoe Municipality is one of the Districts/Municipalities in Volta region and it is located in the central part of the region. It is bounded on the North by Jasikan District, South by Ho – West district, East by Republic of Togo and West by Kpando District. The municipality covers an area of 1172 sq. km. Consisting of 174 communities with a population of 184,743 from the 2010 National Population Census. The population growth rate is 1.9%. There are several institutions in the municipality at various levels on the academic ladder ranging from basic schools to tertiary Institutions. There are 156 Kindergartens, 167 Primary schools, 104 Junior High Schools, 14 Senior High Schools, 6 Technical/Vocational Schools, 2 Colleges of Education, 1 Midwifery College and a satellite campus of University of Health and Allied Science (UHAS). There are about 40% farmers, 35% traders and 25% government workers.

Study design

This study was conducted in 2016, using data from students in various colleges of education including males and females. The preliminary data on subjects such as gender, age, height and weight were gathered, which enable us to calculate the BMI of subjects. Other variables like eating habits, kind of food eaten and physical activities were also obtained. Subjects were captured from December, 2015 to Novermber, 2016.

Data collection

Data was retrieved on subjects with weight problems through an interview. A primary data was obtained from the college students in order to gather first hand information that would reflect the problem among the students.

Data analysis

The data collected was entered into SPSS and analyzed anonymously. A descriptive statistical analysis was done using a Chi-square test which was computed for independent variables (IV) using a 2 × 2 contingency table. Chi-square which is originally known as the Pearson's Chi-square, according to Howell (2011) is referred to both statistical distribution and to a hypothesis testing procedure that produces a statistic test that is approximately distributed as the Chi-square distribution. This test again serves as a "goodness-of-fit" test, where the data are categorized along one dimension, which is commonly known as "contingency table". The categorization is across two or more dimensions. The expected frequencies are computed as $(R_j \times C_k)/N$ where R_j and C_k represent as row and column marginal totals, respectively and *N* is the grand total. The standard Pearson Chi-square statistic is defined as $x^2 = \sum\sum \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$, where *i* and *j* are indexes in the row and column of the contingency table. The resulting test statistics is approximately distributed as x^2 on $(r-1)(c-1)$ degrees of freedom. The likelihood ratio Chi-square builds on the likelihood of a data under null hypothesis relative to maximum likelihood. This is defined as $G^2 = 2\sum O_{ij} \log \frac{O_{ij}}{E_{ij}}$. One of the advantages of the likelihood ratio when using Chi-square is that the $H2$ for large dimensional matrix can be nicely decomposed into smaller components (matrices).

The logistic regression model

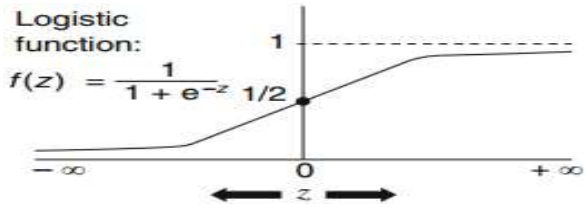
To explain the popularity of logistic regression, the logistic function was shown here, which describes the mathematical form on which

Table 1. The frequency distribution of variables.

Variables	Items	Frequency (%)
Weight outcome	Normal weight	164(53.2)
	Abnormal weight	144(46.8)
Gender	Female	241(78.2)
	Male	67(21.8)
Age	Below 18	50(16.2)
	18 – 29	189(61.4)
	30 – 40	69(22.4)
Eating habit	Below 3 times daily	89(28.9)
	3 Times daily	99(32.1)
	More than 3 times daily	120(39.0)
Physical exercise	No	93(30.2)
	Yes	215(69.8)
Location	Urban	271(88.0)
	Rural	37(12.0)

the logistic model is based. This function, called *f*(*z*), is given by 1 over 1 plus e to the minus *z*. The values of this function was plotted as *z*

varies from $-\infty$ to $+\infty$.



The fact that the logistic function *f*(*z*) ranges between 0 and 1 is the primary reason the logistic model is so popular. The model is designed to describe a probability, which is always some number between 0 and 1. In epidemiologic terms, such a probability gives the risk of an individual getting a disease. The logistic model, therefore, is set up to ensure that whatever estimate of risk we get, it will always be some number between 0 and 1. Thus, for the logistic model, we can never get a risk estimate either above 1 or below 0. This is not always true for other possible models, which is why the logistic model is often the first choice when a probability is to be estimated. The following equations are used to guide and give clear understanding on the logistic model. To obtain the logistic model from logistic, we need to defined the *z* as a linear sum as:

$$z = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \quad (1)$$

Having denied the *z*, we can now write the model as a function of *z*,

$$f(z) = \frac{1}{1 + e^{-z}} \quad (2)$$

We now substitute the linear sum expression for *z* to obtain

$$f(z) = \frac{1}{1 + e^{(\alpha + \sum \beta_i X_i)}} \quad (3)$$

Where, α and β are coefficient of the independent variable *X*, this can be expressed mathematically in more epidemiological term as $P(D = 1|X_1, X_2, \ldots\ldots, X_n)$.

RESULTS

The results of this study are presented in two folds. The first part explained the descriptive statistics of the data analyzed, this includes the percentage distribution of the variables in and cross tabulation of variables like gender and age, while the second part described the logistic model which gives the inferential statistics of the data and determined the predictors of BMI.

It can be observed from Table 1 that, the dependent variable weight has two outcomes namely normal and abnormal weight. The abnormal weight constitutes under-weight, overweight and obese. Out of 308 subjects, 164 (53.2%) had normal weight, while 144 (46.8%) had abnormal weight. In all, five independent variables (gender, age, eating habit, physical exercise and location) were considered, of this, 241 (78.2%) were females and 67 (21.8%) were males, subjects under 18 years were 50 (16.2%), 189 (61.4%) of them were 18 – 29 years and 69 (22.4%) were within ages of 30 to 40 years. Similarly, 89

Table 2. Cross tabulation of gender and weight outcome.

Parameter			Gender		Total
			Male	Female	
Weight outcome	Normal weight	Count	45	119	164
		Expected count	35.7	128.3	164.0
	Abnormal weight	Count	22	122	144
		Expected count	31.3	112.7	144.0
Total		Count	67	241	308
		Expected count	67.0	241.0	308.0

Table 3. Chi-square test.

Paramater	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.662	1	0.010
N of Valid Cases	308		

Table 4. Binary logistic regression predictors of weight outcome.

Variables	B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
							Lower	Upper
Gender	1.241	0.504	6.063	1	**0.014**	3.460	1.288	9.292
Age	1.436	0.337	18.197	1	**0.000**	4.204	2.173	8.132
Eating	-0.610	0.265	5.312	1	**0.021**	0.543	0.323	0.913
Location	-0.128	0.362	0.126	1	0.723	0.879	0.432	1.789
Physical Exercise	1.921	0.554	12.007	1	**0.001**	6.826	2.303	20.232
Constant	-2.123	0.476	19.869	1	0.000	0.120		

(28.9%) had the habit of eating less than three times daily, 99 (32.1%) had the habit of eating three times daily and 120 (39.0%) had the habit of eating more than three times daily. 93 (30.2%) said they had no physical exercise for about a month, while 215 (69.8%) responded yes, meaning they had physical exercise in a month, 271 (88.0%) were residence in the cities, while 37 (12.0%) were residence in villages.

Below is the cross tabulation using a 2×2 contigency table and Chi – square test was computed to determine the association of independent variable (gender) of the dependents variable (weight).

From Table 2 where a cross tabulation of gender and weight outcome was calculated, a total of 67 males and 241 females can be seen. This indicates that more females have a weight problem than males. Table 3 computed the Chi-square test depending on the outcome of Table 2 and the Pearson Chi-square values was found to be $x^2 = 6.662$ with signicance value of 0.01 < 0.05. The study revealed that since the signicance value is less that the alpha value then gender is associated with a weight problem with regard to sample used for the study. That is considering the Hohoe municipality.

Table 4 present the result for the possible predictors of weight outcome for the IV's. The logist model was obtained as:

$$\text{Logit}_{[P(D=1)]} = -2.123 + 1.241\text{Gender} - 1.436\text{Age} - 0.610\text{EH} + 1.921\text{PE}$$

Where, EH is eating habi and PE is physical exercise.

Testing the model, for example, A male who is 18 years, eats 3 times a day and do not do any physical exercise. From Equation 3, we can determine the risk level of subjects having a weight problem. This is computed as:

$$logitP[P(D=1) = \frac{1}{1+e^{-[-2.123+1.241(\text{gender})-1.436(\text{age})-0.601(\text{EH})+1.921(\text{noPE})]}}$$

Table 5. Hosmer and Lemeshow tests.

Chi-square	df	Sig.
6.161	5	.291

Table 6. Model summary.

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
389.923	0.110	0.146

$$logitP[P(D=1) = \frac{1}{1+e^{-[-2.123+1.241(0)-1.436(0)-0.601(1)+1.921(0)]}}$$

$$logitP[P(D=1) = 0.061 \cong 6\%$$

For example, A female who is 18 years, eats 3 times a day and do not take part in any physical exercise.

$$logitP[P(D=1) = \frac{1}{1+e^{-[-2.123+1.241(\text{female})-1.436(\text{age})-0.601(\text{EH})+1.921(\text{no PE})]}}$$

$$logitP[P(D=1) = \frac{1}{1+e^{-[-2.123+1.241(1)-1.436(0)-0.601(1)+1.921(0)]}}$$

$$logitP[P(D=1) = 0.184 \cong 18\%$$

The epidemiological analysis of controlling predictors established that, comparing the two examples, females within the age of 18 who had the habit of eating 3 times daily and do not take part in any physical exercise were three times at higher risk of having a weight problem than their male counterparts who have an estimated risk of only 6% as compared to females with 18% risk.

The Hosmer and Lemeshow test was computed to assess the model fitting. The null hypothesis were set as hypothesized model that fits the data while an alternate hypothesis were set as hypothesized model does not fit the data. The result from Table 5 indicates that, Chi-square value of 6.161 with degree of freedom of 5 has a significance value of 0.291 which is far more that 5% significance level, hence it is concluded that the null hypothesis was rejected or we failed to accept the null hypothesis. With regard to the analysis using epidemiological description, we can say that there was enough evidence to show that the hypothesized model fits the data set used in predicting variables that are associated with weight problems of college students. This therefore maintains that the overall model fit is good. The study further presents the summary of the model where the logs of the likelihood and variances level of prediction are done.

It was identified and deduced from Table 6 that Pseudo R – square stated 14.6% of the variance in predicting whether or not students have a weight problem considering the following variables: gender, age eating habit, physical exercise and location (place of residence). Classification table was also computed to determine the overall percentage predicted outcome of a weight problem and its associated contributing factors among college students.

The classification result of predicted percentage is shown in Table 7. It displays that about 80.5% could be predicted as normal weight, while 43.1% could be predicted as abnormal weight. It is also proper to establish that 63.0% of the overall subjects were correctly classified.

DISCUSSION

The findings of the study are identified and presented in two folds. First of all, a descriptive discussion was done followed by the presentation of the logistic model. It should be noted that there were 164 subjects with normal weight, while 144 of them had abnormal weight (weight problem) and when considering gender, there were more females in the study that their male counterparts. The ages of subject's ranges from 17 to 38 and it was categorized into three groups. 16.2% of the subjects were

Table 7. Classification table.

Observed		Predicted		
		Weight outcome		Percentage correct
		Normal weight	Abnormal weight	
Weight outcome	Normal weight	132	32	80.5
	Abnormal weight	82	62	43.1
Overall percentage				63.0

under 18 years, 22.4% were within 30 to 40 years, while 61.4% constituting majority of the subjects were within 18 to 29 years. In like manner, 28.9% create a habit of eating less than 3 times daily, 32.1% eats 3 times daily and 39.0% eat more than 3 times daily. It was also maintained that 30.2% were not involved in physical exercise and as many as 69.8% involved themselves in physical exercise and the majority of them were residing in the towns and cities, while only 12.0% were in the rural communities.

Cross tabulation of 2 by 2 was computed for gender as a cofounder to check the test of association of weight problem. It was deduced from the test that a Chi–square of 6.662 was computed with a degree of freedom of 1 showing a significance value of 0.01 which less than 5% significance alpha value, therefore gender is associated with weight outcome holding for other variables in the test. This was followed by the logistic model table which presents all the IV's tests for the DV's in the study. It was indicated that gender, age, eating habits and physical activities were variable found in the model and only location (place of residence) was not found in the model. It implies that the IV's found in the model are contributing factors to the DV's and that location was not a possible predictor of weight problem.

Conclusion and recommendation

The findings maintained that although there might be other predictors surrounding weight outcome, four (4) variables out of five in this study were proven to be the possible factors associated with weight outcome among the subjects sampled for the purpose of this study. These predictors are gender, age, eating habits and physical exercise. An epidemiological example reveals that a female who is 18 years, that eats normally and do not do any physical exercise is at a higher risk of weight problem as compared to the male counterpart. Furtheremore, theoretical factors according to Strong et al. (2008) have been reported as influential in exploratory health behaviors and behavior change among young adults. In addition, Ha and Caine-Bish (2009) reported a higher incidence in college students adopting unhealthy eating behaviors such as skipping meals, frequent snacking on calorie-dense food, and engaging in unhealthy weight-loss or weight-gain methods. It is therefore recommended that, since gender, age, eating habits and physical exercise were part of the model, thus significant to this study, every individual must be watchful about his/her weight outcome depending on the number of times they eat without doing any exercise. Moreover, the model was tested and shows that females are more vulnerable as compared to their male counterpart considering age.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interest.

REFERENCES

Ha EJ, Caine-Bish N (2009). Effect of nutrition intervention using a general nutrition course for promoting fruit and vegetable consumption among college students. Journal of Nutrition Education and Behavior 41(2):103-109.

Howell DC (2011). Chi-square test: analysis of contingency tables. In International encyclopedia of statistical science. Springer Berlin Heidelberg pp. 250-252.

Keski-Rahkonen A, Kaprio J, Rissanen A, Virkkunen M, Rose RJ (2003). Breakfast skipping and health-compromising behaviors in adolescents and adults. European Journal of clinical nutrition 57(7):842-853.

Khawaja NG, Dempsey J (2008). A comparison of international and domestic tertiary students in Australia. Australian Journal of Guidance and Counselling 18:30-46.

Laquatra I (2004). Nutrition for weight management. In Krause's Food, Nutrition, and Diet Therapy. Ed. by Mahan LK, Escott-Stumps S, 11thed. Saunders Company,

Popkin BM (1994). The nutrition transition in low-income countries: an emerging crisis. 52(9):285-98.

Rahman A, Hakim MA (2017). Modeling health status using the logarithmic biophysical modulator. Journal of Public Health and Epidemiology 9(5):145-150.

Rahman A, Hakim MA (2016). Measuring modified mass energy equivalence in nutritional epidemiology: a proposal to adapt the biophysical modeling approach. International Journal of Statistics in Medical Research 5(3):219-223.

Strong K, Parks S, Anderson E, Winett R and Davy B. (2008). Weight gain prevention: identifying theory-based targets for health behavior change in young adults. Journal of The American Dietetic Association 108(10):1708-1715.

Swinburn BA, Caterson I, Seidell JC, James WPT (2004). Diet, nutrition and the prevention of excess weight gain and obesity. Public Health Nutrition 7(1A):123-146

World Food Programme (2008). https://www.wfp.org/stories/high-food-prices-why-different-2008

World Health Organization (WHO) (2013). World health Statistics, a wealth of information on global public health. *http://apps.who.int/iris/handle/10665/82058*

WHO (2000). Obesity: preventing and managing the global epidemic (No. 894). World Health Organization.