Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.7, No.7, 2017

Log-linear Model of Diabetic Patients of Plateau State General Hospitals

Longben, Philip Gobum[1] ; Dusu, Paul Bot [2*]; Peter, Shut Audu [4] ; Selchak, Selbut D [3]; Nanbol, Helen D. [3]; Gotan, Anthony [2]

1. School of Health Information Management and Biostatistics, Plateau State College of Health Technology Pankshin, PMB 13, Plateau State-Nigeria.
2. General Studies Department, Plateau State College of Health Technology Pankshin, PMB 13, Plateau State-Nigeria.
3. School of Medical Laboratory Science, Plateau State College of Health Technology Pankshin, PMB 13, Plateau State-Nigeria.
4. School of Community Health, Plateau State College of Health Technology Pankshin, PMB 13, Plateau State-Nigeria.

* dusupaul@yahoo.com & pauldusu@gmail.com

Abstract

This study investigates Diabetes Mellitus, which is a disorder that is assuming pandemic proportion worldwide. With secondary data from Plateau State General Hospitals and a log-linear model analysis carried out using statistical package (SPSS version 21), to find out if there is a relationship between diabetes, gender and Body mass index (BMI). The result shows that there is a significant relationship between female as reported to BMI and diabetes with a P- value of 0.0007, and male as reported to BMI and diabetes with a P- value of 0.000. This i-ndicates an interaction between gender, BMI and diabetes because all their P-values are less than the alpha level of 0.05. The study recommends public education on regular BMI checking, regular exercise and proper diet by both sex to avoid obesity which is generally associated with an increase in risk of diabetes mellitus.

Keywords: Diabetes Mellitus, Obesity, Body mass index and log-linear analysis.

1. Introduction

Diabetes Mellitus is a chronic disorder that is not only assuming pandemic proportions worldwide but also poised to affect the developing countries of the world much more than their developed counterparts. Kahn (2005) pointed out that there is compelling data to show an increase incidence and prevalence of diabetes mellitus on the continent of Africa with 1% in the rural areas and range from 5% to 7% in urban sub-Sahara Africa. This figures must have increase since diabetes mellitus has been described as an uncommon disorder in Africa by scholars, among them is Colagiuri, Crystal, Colagiuri, Magliano, Jonathan, Paul and Ian (2010). According to Ogbera (2014) Nigeria with a population of 160 million people accounts for one-sixth of Africans diabetes mellitus cases. Being a non-communicable disease, diabetes patients suffer a reduction in life expectancy and quality of life. Diabetes is now cited as a global epidemic and intertwined with obesity epidemic. The international Diabetes Federation estimates that there were 184 million people with diabetes in 2003 and predicts an increase to 324 million in 2025. The world health organization (WHO) estimates an increase from

171 million in 2000 to 366 million in 2030 with approximately 70% growth is predicted to occur in the developing world and will increasingly affect people aged younger than 65 years who are still in the productive stages of their life cycle. The increase in the younger age group being diagnosed with diabetes mellitus type 2 pose an economic threat over and above the more direct disease cost to the public as pointed out by (Colaguiri et al, 2010).

An increase in body fats is generally associated with an increase in risk of metabolic disease such as diabetes mellitus. Metabolic syndrome (Mets) and body mass index (BMI) are established independent risk factors in the development of diabetes. Increase BMI has been associated with metabolic and cardiovascular risk factor of diabetes but there is increasing evidence that sub-phenotype of obesity exist that appear to deviate from the standard dose-response relationship between increase BMI and its adverse clinic outcomes. Normal weight is associated with three to four folds higher risks for diabetes as compared with control subjects, while metabolically health but obese individual have been identified, who despite having BMI exceeding $30kgm^{-2}$ are relatively insulin sensitive and having a rather favourable cardiovascular risk people with a three to four fold lower risk for diabetes as compared with obese insulin resistance individual (Malcolim, 2015). Diabetes creates problem with hormone insulin. Generally, the pancreas releases insulin to assist the body store and use the sugar and fat from the food that we eat. Diabetes mellitus is a collection of metabolic disorder owing to relative insulin deficiency or resistance or both. Diabetes affects the body ability as well as capacity to use blood sugar energy.

Farzard, Mohamadreza, Maryam, Davood and Fercidoun (2011) opined that since no population-base study prospectively has examined the joint relationship between BMI with diabetes, an unanswered question remains to be whether the impact of diagnosis of different obesity phenotype on prediction of incident diabetes differs by body weight. For this, in the era of globalization, an attention is given to concentrate to build mathematical model for the diabetic patient, hence this research will provide a statistical model (Log-linear model) using data based on patients' attendance at a hospital to investigate the combine relationship of BMI and diabetes to understand if diabetes modifies increase effect of BMI on diabetes.

Review of literatures by Adarabioyo (2014) on Child Mortality, Okoli, Oneyeagu, and Osuji, (2014) on Testing the Significances of Interaction of Three Categorical Variables, Rafiqui (2012) on distribution of diabetes patient for both sexes associated with age and Sharmin (2014) on Examining Potential Risk Factor to Acute Pancreatitis Disease; prompted the use of log-linear as a good statistical tool for this work.

1.1 Conceptual Definition of Terms

Diabetes Mellitus: Colagiuri et al (2010); Farzad et al (2011) and Kahn (2005) highlighted that, Diabetes is a disease that create problem with hormone insulin. Generally, the pancreas releases insulin to assist our body store and use the sugar and fat from the food that we eat. Diabetes also familiar as diabetes mellitus is a collection of metabolic disorder characterized by chronic hyper glycaemia owing to relative insulin deficiency or resistance or both. Diabetes affects the body ability as well as capacity to use blood sugar energy. Diabetes affects people of all kinds of socio-economic conditions all over the world.

Body Mass Index: Eknoyon (2007) and Malcolin (2015) view body mass index as a value derived from the mass (weight) and defined as the body mass divided by the square of the body height, and is universally expressed in unit kg/m^2, resulting from mass in kilograms and height in meters. Commonly accepted BMI ranges are underweight under 18.5, normal 18.5 to 25, overweight 25 to 30 obese over 30 and is commonly employed among children and adults to predict health outcomes. The BMI trait is influenced by both genetic and non-genetic factors and it provides a paradigm to understand and estimate the risk factors for health problems. BMI provides a simple numeric measure of a person's thickness or thinness, allowing health professionals to discuss weight problems more objectively with their patients.

Log-Linear Analysis: Adarabioyo (2010); Feild (2005); Jeansonne (2010) and Okoli et'al (2014) contributed that, Log linear analysis is a technique used in statistics to examine the relationship between more than two categorical variables. The technique is used for both hypothesis testing and model building. In both these uses, models are tested to find the most parsimonious (that is least complex) model that best accounts for the variance in the observed frequencies. Log linear models was primarily developed during the 1960s and first comprehensively described the methodology for the general statistical community. By linear model methodology is most appropriate when there is no clear distinction between the response variable and explanatory variable. The log linear model point of view treats all variables as response variables and the focus is on statistical independence and dependence. Log linear modelling of categorical data is analogous to correlation analysis for normally distributed response variables and is useful in assessing patterns of statistical dependence among subjects of variables. The model fits well when the residuals (observed-frequencies) are closer to the observed frequencies (that is if the likelihood ratio chi-square statistic is non-significant) but if the likelihood ratio chi-square statistics is significant the model does not fit well (that is calculated expected frequencies are not close to observed frequencies).

Obesity: Colagiuri et al (2010), Malcolim (2015) and WHO (2000) are of the view that obesity is a condition where a person has accumulated so much body fat that it have a negative effect on their health and its' one of the most pervasive chronic diseases in need of new strategies for medical treatment and prevention. If a person's bodyweight is at least 20% higher than it should be, such person is obese. A body mass index of between 25 and 29.9 is considered obese.

2. Objective of the Study

The objective of this study is to apply Log-linear model on body mass index and Diabetes data by;

i) Examining the relationship between body mass index and Diabetes.

ii) Evaluating and comparing the association between body weight change and the risk of Diabetes Mellitus.

3. Hypothesis

H_O: There is no significant relationship between Gender as regard to Body Mass Index and Diabetes.

H_1: There is significant relationship between Gender as regard to Body Mass Index and Diabetes.

4. Methodology

4.1 Source of Data

Secondary data from six Plateau State General Hospitals which are Pankshin General Hospital, Dengi General Hospital, Mangu General Hospital, Langtang General Hospital, Shendam General Hospital and Barkin ladi General Hospital was used in this research.

4.2 Data Analysis Technique

The collected secondary data was analysis by use of a Log-linear model.

4.2.1 Log-linear model

Model components for this work main effects and interactions are the relationship between three variables (Variable A, variable B, and variable C) that produce seven model components in the saturated model which are the three main affects (A, B, C), the three two-way interactions (AB, AC, BC) and the one three-way interactions (A B C).

In this model observed frequencies equal expected frequencies, therefore in the likelihood ratio chi-square statistic, the ratio is $\frac{O_{ij}}{\Sigma_{ij}}$ =l and ln (l)= 0. This results in the likelihood ratio chi-square statistic being equal to 0, which is the best model fit. With the three variables (A,B,C) the saturated model has the following log-linear equation;

$$In(f_{ijk}) = \lambda + \lambda^{A} + \lambda^{B} + \lambda^{C} + \lambda^{AB} + \lambda^{AC} + \lambda^{BC} + \lambda^{ABC}$$

where, F_{ijk} = expected frequency in cell ijk,

λ = the relative weight for each variable (Agresti, 2007).

The following models refer to the traditional chi-square test where two variables, each with two levels (2x2 table) are evaluated to see if an association exist between the variables and is extended from the $2 * 2$ table to the $s * r$ table.

$$In(f_{ij}) = m + \lambda_i^{A} + \lambda_j^{B} + \lambda_{ij}^{AB}$$

$In(f_{ij})$ is the log of the expected cell frequency of the cases for cell ij in the contingency table.

m is the overall mean of the natural log of the expected frequencies.

λ is the terms each represent ‘effect’ which the variables have on the cell frequencies.

i and j, refer to the categories within the variables.

Hence

λ_i^{A} is the main effect for variable A,

λ_j^{B} is the main effect for variable B

λ_{ij}^{AB} is the interaction effect for variables A and B.

The above model is a saturated model because it includes all possible one-way and two effects. Given that the saturated model has the same amount of cells in the contingency table as it does effects the expected cell frequencies will always exactly match the observed frequencies with no degrees of freedom remaining. Thus, in order to find a more parsimonious model that will isolate the effects best demonstrating the data patterns, a non-saturated model will be sought. This is achieved by setting some of the effect parameters to zero.

The non-saturated model fitted the independence model because it lacks an interaction effect parameter between A and B implicitly; this model holds that the variables are unassociated. The independence model is analogous to the chi-square analysis testing the hypothesis of independence.

4.2.2 Fitting Log Linear Models

An interactive proportional fitting algorithm (Deming-Stephen algorithm) is used to generate expected frequencies. This procedure uses marginal tables fitted by the model to insure that the expected frequencies sum across the other variables to equal the corresponding observed marginal tables (Knoke and Burke, 1980)

4.2.3 Log Linear Model for the $s * r$ Table

A sample of n observations is classified with respect to two categorical variables, one having s levels and the other having r levels, the resulting frequencies are shown in the table below (Cell counts in an $s * r$ contingency table) according to Maura *et al* (2000).

Table 1: Cell counts in an $s * r$ contingency table

LEVEL OF X	LEVEL OF Y				TOTAL
	1	2 …	r		
1	n_{11}	n_{12}…	n_1r		n_{1+}
2	n_{21}	n_{22}…	n_2r		n_{2+}
.	.	.		.	.
.	.	.		.	.
.	.	.		.	.
S	n_{s1}	n_{s2}…n_{sr}			n_{st}
TOTAL	n_{+1}	n_{+2}…	n_{+r}		N

The saturated model is $\log(m_{ij}) = \mu + \lambda_i^x + \lambda_j^x + \lambda_{ji}^y + \lambda_{ij}^{xy}$.

i = l, ……., s, j= 1…………, r

where $m_{ij} = n \prod ij$ is the expected frequency in the (i, j) cell.

The parameter m is fixed by the sample size n and the model has s + r + sr parameters, λ_i^x, λ_j^y *and* λ_{ij}^{xy} they sum-to zero constraints.

$$\sum_{1=1}^{s} \lambda_i^x = 0, \ \sum_{j=1}^{r} \lambda_j^y = 0, \ \sum_{i=1}^{s} \lambda_{ij}^{xy} = 0, \ \sum_{j=1}^{r} \lambda_{ij}^{xy} = 0.$$

This implies (s-1) + (r-1) + (s-1) (r-1) = sr-1 parameters and zero df for testing lack of fit.

Letting $m_{ij} = n_i + n + j/n$, the likelihood ratio statistic.

$$G^2 = 2 \sum_{i=1}^{s} \sum_{j=1}^{r} n_{ij} \log \left({}^{n_{ij}}/m_{ij} \right).$$

Tests the null hypothesis $H_0: \lambda_{ij}^{xy} = 0$, for i = 1,……, s-1, j=1,….., r-1 under the null hypothesis of independence, G^2 has an approximate chi-square distribution with (s-1) (r-1) df.

If H_o is true, the reduced model $\log(m_{ij}) = m + \lambda_i^x + \lambda_j^y$ is the model of independence of x and y.

This model has (s-1) + (r-1) linearly independent λ parameters and (s-1) (r-1) df, for testing lack of fit.

4.3 Statistical Analysis

The statistical analysis (Log-linear model) considered a three-way order association, which are Gender, BMI and Diabetes Mellitus.

The tables below show the data obtained for reported cases of Diabetes according to the three-way order association earlier mention.

Table 1: Diabetes Data for male and female

		DIABETES		
GENDER	**BODY MASS INDEX**	**YES**	**NO**	**TOTAL**
MALE	LOW	176	35	211
	NORMAL	77	44	121
	HIGH	251	39	290
	TOTAL	**504**	**119**	**623**
FEMALE	LOW	219	26	245
	NORMAL	159	36	195
	HIGH	381	40	421
	TOTAL	**759**	**102**	**861**

Source: *Plateau State General Hospitals (Longben, 2016)*

In order to establish the association between Gender, Body Mass Index and Diabetes, Chi-square test is run in this format; Gender-female, BMI and Diabetes keeping Gender-male = control

Table 2: Chi-Square Test for Female

GENDER		Value	df	Asymp. Sig. (2-sided)
FEMALE	Pearson Chi-Square	10.746[a]	2	0.005
	Likelihood Ratio	9.865	2	0.007
	Linear-by-Linear Association	0.712	1	0.399
	N of Valid Cases	861		

Table-2 shows that; Chi-Square = 10.746, df = 2 and P = 0.005, which is significant since the P-value at 0.005 is less than the alpha value of 0.05. The same goes to G-statistic which is the likelihood ratio with a P-value of 0.007 which is also less than alpha value of 0.05. Therefore we reject the null hypothesis and conclude that there is a significant relationship between female as regard to body mass index and diabetes.

Table 3: Chi-Square Tests for Male

GENDER		Value	Df	Asymp. Sig. (2-sided)
MALE	Pearson Chi-Square	30.345[a]	2	0.000
	Likelihood Ratio	27.117	2	0.000
	Linear-by-Linear Association	1.597	1	0.206
	N of Valid Cases	622		

Table-3 shows that: Chi-square value is 30.34 with df = 2 and P = 0.000 which indicate significant difference among the males.

Which is also significant since the P-value at 0.000 is less than the alpha value of 0.05. The same goes to G-statistic which is the likelihood ratio with a P-value of 0.000 which is also less than alpha value of 0.05. Therefore we reject the null hypothesis and conclude that there is a significant relationship between male as regard to body mass index and diabetes.

Table 4: Log-linear result

Effect	Df	Partial Chi-Square	Sig.	Number of Iterations
GENDER*BMI	2	7.798	0.020	2
GENDER*DIABETES	1	20.466	0.000	2
BMI*DIABETES	2	14.751	0.001	2
GENDER	1	25.530	0.000	2
BMI	2	641.197	0.000	2
DIABETES	1	336.237	0.000	2

Table-4 shows that there is an interaction between the gender and BMI, gender and diabetes and finally BMI and diabetes, since all their p-values are less than the alpha level of significant which is 0.05.

4.4 Discussion

The results in tables 2 and 3, infer that there is an association between BMI and Diabetes between the male and female P-values of 0.005 and 0.000 respectively which is less than the alpha value of 0.05. Hence, the null hypothesis was rejected and the conclusion was that there exist a relationship between body mass index and diabetes.

4.5 Conclusion

The results of the study indicated high explanatory power for both gender, body mass index and diabetes. This suggest that increase in body mass index or overweight increase the risk of diabetes mellitus for both male and female as pointed out by our result. Furthermore, the results for male regards to body mass index and diabetes also shows that there is a significant association between male and diabetes. The male are more prone to diabetes a little than their female counterpart.

4.6 Recommendation

This work recommends that; An increase in body mass index is generally associated with an increase in the risk of diabetes mellitus such as type 2 diabetes mellitus therefore the general public is advised to watch their weight to avoid the risk of diabetes mellitus.

The general public should be educated on how to check their weight at regular intervals in order to avoid obesity, be educated on the kind of diet they eat and be informed that regular exercise can help in reducing the rate of diabetes mellitus.

References

Adarabioyo, M .I, (2010). Application of Log Linear Model to determinants of Child mortality in Nigeria. A paper presented to the Department of Mathematics and Physical Science AfeBabalola University Ado-Ekiti.

Agresti, A. (2007). An Introduction to Categorical Data Analysis (2nd ed.). Hoboken, N Wiley Inter-science. P.212. ISBN 978-0-22618-5.

Colagiuri, S., Crystal, M.Y.L., Colagiuri, R., Magliano, D., Jonathan, E.S., Paul, Z.Z. and Ian, D.C. (2010). The cost of overweight and obesity in Australia. *The Medical Journal of Australia*; 192 (5): 260-264.

Ekoyon, J. (2007). Body mass index to determine whether people are healthy. Msc. Thesis.

Ferzad, H. Mohamadreza, B., Maryam, S., Davood, K. and Fercidoun, A. (2011). Diabetes prediction, lipid accumulation product, and adiposity, lipid and glucose study.

Field, A. (2005). Discovering Statistics using SPSS (2nd edition). Thousand Oaks, CA: Sage Publications. pp. 695-718.

Jeansonne, A. (2010). Log Linear Models of Categorical Data. International Journal of statistical Association. 18(11); 512-569.

Kahn, R. (2005). American Diabetes Association; European Association for the study of Diabetes. The Metabolic Syndrome: time for a critical Appraisal; joint statement from the American Diabetes Association and the European Association for the study of Diabetes. Diabetes Care.28:2289-304

Malcolim, J. (2015). Body mass index in Adults (BMI calculator for Adults). American Heart Association.

Ogbera, A.O. (2014). Diabetes Mellitus in Nigeria the past, present and future. World Journal of Diabetes.

Okoli, C. Oneyeagu, N. and Osuji O. (2014). Loglinear Models Selection for Contingency Tables. Journal of National Science Research Information Criteria.

Longben, Philip Gobum. (2016). Log-Linear Modelling of Diabetic Patients. A study of Plateau State General Hospitals, Postgraduate Diploma in Statistics Thesis, Department of Mathematics/Statistics and Computer Science, University of Agriculture, Makurdi. Benue State, Nigeria (unpublished).

Rafiqui, I. (2012). Modelling of Diabetes Associated with Age Polynormial Approach, International Journal of Statistics and Application. Vol. 1 No.1 pp.1-5

Sharmin, I. Mustafa K. Zamalia M. Saperi, S. and Niswah, N. A. (2014). Examining Potential Risk Factors to Acute Pancreatitis Disease.

World Health Organization (2000). Obesity: Preventing and Managing the Global Epidemic. Report of a WHO Consultation. World Health Organ Tech Rep Ser. 894:i-xii,1-253.