Websites:
http://www.nbmedicine.org
http://www.sciencepub.net/nurse

Emails:
editor@sciencepub.net
nbmeditor@gmail.com

Biomedicine and Nursing

MARSLAND PRESS
Multidisciplinary Academic Journal Publisher

Logistic Regression Analysis on the Diabetic Status of Patient.

Oladimeji O. A.(1), Ikotun D. O.(2) and Ogunbanwo T. S. (3)

[1]Department of Statistics, Federal Polytechnic, Ile-Oluji, Ondo State, Nigeria
(adedipupo.oladimeji@gmail.com)
[2]Department of Statistics, Federal Polytechnic, Ile-Oluji, Ondo State, Nigeria
(fisikotun@gmail.com)
[3]Department of Mathematics and Statistics, Osun State College of Technology, Esa-Oke, Osun State, Nigeria
(tolualone@gmail.com)

ABSTRACT: Diabetes is an epidemic disease suffered by people all over the world. The result of this ailment is long time pain an eventual death if not properly taken care of. The associated factors that causes this ailment are high blood pressure, family history, alcohol consumption and so on. The data used for this paper was a secondary data collected from Ladoke Akintola University Teaching Hospital (LAUTECH), Ogbomoso, Oyo state, Nigeria. The Methodology was Binary Logistic Regression. From the Binary Logistic Regression result, the Wald Test reveals that all the predictor variables except high cholesterol level were significant at 5 percent level of significance. The likelihood ratio test revealed that at-least one of the predictor variables contribute significantly to the model at 5 percent level of significance. The findings of the paper recommend that Access to use of health services need to be increased in our communities particularly for diabetes in order to make individuals to visit the place early for treatment than for the disease to reach its climax.

[Oladimeji O. A., Ikotun D. O. and Ogunbanwo T. S. **Logistic Regression Analysis on the Diabetic Status of Patient.** *Biomedicine and Nursing* 2021;7(4):21-30]. ISSN 2379-8211 (print); ISSN 2379-8203 (online). http://www.nbmedicine.org. 3. doi:10.7537/marsbnj070421.03.

Keywords: Diabetes, Binary Logistic, Smoking, Cholesterol, Hypertensive Status

INTRODUCTION

The term *diabetes*, without qualification, usually refers to diabetes mellitus, which roughly translates to excessive sweet urine (known as "glycosuria"). Several rare conditions are also named diabetes. The most common of these is diabetes insipidus in which large amounts of urine are produced (polyuria), which is not sweet (insipidus meaning "without taste" in Latin).

Criteria for the diagnosis and classification of diabetes have been revised several times and can differ between countries. The current WHO classification (2003), is as follows:

- type 1 (pancreatic beta-cell destruction leading to absolute insulin deficiency),
- type 2 (insulin resistance and relative insulin deficiency)

other specific types of diabetes
gestational diabetes

Type 1 diabetes mellitus is characterized by loss of the insulin-producing beta cells of the islets of Langerhans in the pancreas leading to insulin deficiency. This type of diabetes can be further classified as immune-mediated or idiopathic. The majority of type 1 diabetes is of the immune-mediated nature, where beta cell loss is a T-cell mediated autoimmune attack. There is no known preventive measure against type 1 diabetes, which causes approximately 10% of diabetes mellitus cases in North America and Europe. Most affected people are otherwise healthy and of a healthy weight when onset occurs. Sensitivity and responsiveness to insulin are usually normal, especially in the early stages. Type 1 diabetes can affect children or adults but was traditionally termed "juvenile diabetes" because it represents a majority of the diabetes cases in children.

Type 2 diabetes mellitus is characterized by insulin resistance which may be combined with relatively reduced insulin secretion. The defective responsiveness of body tissues to insulin is believed to involve the insulin receptor. However, the specific defects are not known. Diabetes mellitus due to a known defect are classified separately. Type 2 diabetes is the most common type.

In the early stage of type 2 diabetes, the predominant abnormality is reduced insulin sensitivity. At this stage hyperglycemia can be reversed by a variety of measures and medications that improve insulin sensitivity or reduce glucose production by the liver.

Gestational diabetes mellitus (GDM) resembles type 2 diabetes in several respects, involving a combination of relatively inadequate insulin secretion and responsiveness. It occurs in about 2%–5% of all pregnancies and may improve or disappear after delivery. Gestational diabetes is fully treatable but requires careful medical supervision throughout the pregnancy. About 20%–50% of affected women develop type 2 diabetes later in life.

The following tests are used for clinical diagnosis of diabetes:

- Afasting plasma glucose (FPG) test measures blood glucose in a person who has not eaten anything for at least 8 hours. This test is used to detect diabetes and pre-diabetes.
- Anoral glucose tolerance test (OGTT) measures blood glucose after a person fasts at least 8 hours and 2 hours after the person drinks a glucose-containing beverage. This test can be used to diagnose diabetes and pre-diabetes.
- Arandom plasma glucose test, also called a casual plasma glucose test, measures blood glucose without regard to when the person being tested last ate. This test, along with an assessment of symptoms, is used to diagnose diabetes but not pre-diabetes.

Test results indicating that a person has diabetes should be confirmed with a second test on a different day.

Signs and symptoms

The classical symptoms of diabetes are polyuria (frequent urination), polydipsia (increased thirst) and polyphagia (increased hunger). Symptoms may develop rapidly (weeks or months) in type 1 diabetes while in type 2 diabetes they usually develop much more slowly and may be subtle or absent. Prolonged high blood glucose causes glucose absorption, which leads to changes in the shape of the lenses of the eyes, resulting in vision changes; sustained sensible glucose control usually returns the lens to its original shape. Blurred vision is a common complaint leading to a diabetes diagnosis; type 1 should always be suspected in cases of rapid vision change, whereas with type 2 change is generally more gradual, but should still be suspected.

People (usually with type 1 diabetes) may also present with diabetic ketoacidosis, a state of metabolic dysregulation characterized by the smell of acetone; a rapid, deep breathing known as Kussmaul breathing; nausea; vomiting and abdominal pain; and an altered states of consciousness. A rarer but equally severe possibility is hyperosmolar nonketotic state, which is more common in type 2 diabetes and is mainly the result of dehydration. Often, the patient has been drinking extreme amounts of sugar-containing drinks, leading to a vicious circle in regard to the water loss.

A number of skin rashes can occur in diabetes that is collectively known as diabetic dermadromes.

Causes

The cause of diabetes depends on the type. Type 2 diabetes is due primarily to lifestyle factors and genetics. Type 1 diabetes is also partly inherited and then triggered by certain infections, with some evidence pointing at Coxsackie B4 virus. There is a genetic element in individual susceptibility t o some of these triggers which has been traced to particular HLA genotypes (i.e., the genetic "self" identifiers relied upon by the immune system). However, even in those who have inherited the susceptibility, type 1 diabetes mellitus seems to require an environmental trigger.

INCIDENCE OF DIABETES IN NIGERIA

Diabetes complications in nigeria Complications of DM (acute and chronic) can be severe, debilitating and fatal. The longer the duration of the illness, the greater the possibility of an end organ complication. Some reports have shown that populations of African origin have high prevalence of microvascular (and low of macrovascular) complications [36-39] partly due to co-existent Hypertension, inappropriate diabetic control and limited access to care. Their reports indicate that 21% – 25% of individuals with type 2 diabetes have retinopathy at diagnosis of diabetes and overall, retinopathy affects 15% - 55%, with a high proportion of proliferative retinopathy and macular edema [36]. In addition to diabetic retinopathy being a leading cause of adult blindness, diabetic subjects are six times more prone to cataracts and 1.4 times more susceptible to open- angle glaucoma [37,38]. In cohorts with mean diabetes duration of 5years – 10 years, 32% - 57% have micro or macroalbuminuria, and a third to half of people on maintenance hemodialysis have diabetes [36]. Coronary heart disease affects 5% - 8% of individuals with type 2 diabetes and cardiomyopathy up to 50% of all patients with type 2 DM. Approximately 15% of people with stroke have diabetes while up to 5% of individuals with diabetes present with Cerebrovascular accidents at diagnosis. Prevalence of peripheral vascular disease (PVD) varies across sites from 4% to 28% and only 20% of diabetic foot lesions are attributable to PVD [39].

Objectives of the study

The aim of this paper is to use the methodology of binary logistic regression on the data on the diabetic status of patients.

i. give a descriptive description of the diabetic status across the cholesterol level ,Glucose level, Hypertensive status, Smoking Status and Family history, Smoking status and hypertensive status.
ii. to fit a binary logistic regression model
iii. to test for the significance of the individual predictors

LITERATURE REVIEW

Diabetes mellitus is defined as a group of metabolic diseases whose common feature is an elevated blood glucose level (hyperglycemia). Chronic hyperglycemia is associated with the long-term consequences of diabetes that include damage and dysfunction of the cardiovascular system, eyes, kidneys and nerves.

Diabetes mellitus (DM) was estimated to be the 29th leading cause of burden of disease in the world in 1990, accounting for 1.1% of total years lived with disability (YLD), around the same percentages respiratory infections or malignant neoplasms.

In the Version 1 estimates for the Global Burden of Disease (GBD) 2000 study, published in the World Health Report 2001, DM is the 20th leading cause of YLDs at global level, accounting for 1.4% of total global YLDsCriteria for the diagnosis and classification of diabetes have been revised several times and can differ between countries. The current WHO classification of disorders of glycaemia is as follows:

i. type 1 (pancreatic beta-cell destruction leading to absolute insulin deficiency),
ii. type 2 (insulin resistance and relative insulin deficiency)

· other specific types of diabetes
· gestational diabetes

The onset of type 1 diabetes is most common in children or young adults and accounts for around 10% or less of the total number of people with diabetes. Type 2 diabetes accounts for almost all of the remaining cases of diabetes as the other forms are rare. Type 2 diabetes is a condition that predominantly affects middle-aged and older people but prevalence is increasing among children and young adults in countries with a high prevalence of obesity.

Diabetes was considered as a problem of developed countries in the past, but now people of developing countries are also at equal risk of diabetes and migrants may often be more likely to develop diabetes than non migrant, Carballo and Frederik (2006).

Also, from the work of Cockram(2000), diabetes was recognized as a major global epidemic, with the prevalence of the disease rapidly increasing in many developed and/or developing Asian countries.

Both diabetes and hypertension are independent risk factors for cardiovascular disease which indicates that the co-existence of these conditions in a patient imposes a need for a significant blood pressure control (135/80mmHg) than the goal blood pressure recommend for patient who does not have diabetes (140/90mmHg).

Apart from the issue of high blood pressure as a risk factor for diabetes, Haffner and Heinrich (2003), also found that the most essential risk factors are obesity weight gain and sedentary life style.

Keil et al. (2001) and Kolenda and Muller (2005), grouped the risk factors of diabetes into non modifiable and lifestyle associated factors respectively.

Increased age, male sex and a positive family history count to the non modifiable Cardiovascular risk factors whiles smoking, overweight and obesity, diabetes mellitus, hypertension, physical inactivity, Hypercholesterolemia as well as hypertriglyceridemia belongs to the lifestyle associated risk factors.

Hippocrates (2005), recognized that sudden death from diabetes is more common in those who are naturally fat than in the lean.

Bray (2004), compared whites and blacks on cholesterol levels. He suggested that control of overweight or high cholesterol level would eliminate 48% of the diabetes in whites and 28% in blacks.

Again, the National Institute for Health (1999), also found in their study that the prevalence of type 2 diabetes has tripled in the last 30 years and much of the increase is related with increase in cholesterol levels. People with Body mass index of 30kg or more have five fold greater risk of diabetes than those with < 25kg BMI.

A study done in Poland was concluded with the following results: The prevalence of diabetes or impaired glucose tolerance was found in 5.3% and 92.8% of subjects having diabetes or impaired glucose tolerance were either obese or have high cholesterol level and 32.4% had hypertension.

Also, according to the research of Mohan et al. (2004) increase in cholesterol level has become the emerging burden of risk factors for non-communicable disease, blood pressure, and is now a major public health problem for all age groups. They said blood pressure is frequently elevated in children with high cholesterol level or increase in fat as compared to lean subjects. This they said is possible related to their sedentary lifestyle, altered eating habits, increased fat content of diets and decreased physical activities.

The last but not the least findings found in Africa revealed the effect of high cholesterol on blood pressure is higher in males than in females (regression coefficient 0.64 and 0.38 respectively). Mufunda et al. (2006).

The last recent study in Ghana has shown that adjusted odds ratio for developing hypertension for overweight or high cholesterol were 5.8 and 8.9 respectively. Addo et al. (2008).

High Blood Pressure/ Hypertension

Blood pressure plays an important part in the management of diabetes. High blood pressure (hypertension) adds to the workload of the heart, arteries and kidneys. Damage to kidneys, eyes and feet are long-term complications that can go along with a diagnosis of diabetes. Other health risks include heart disease and strokes.

Global assumptions are difficult due to heterogeneity between countries. According to a systematic review of studies reporting data from 1980 and 2004, the overall worldwide prevalence of high blood pressure was approximately 26% in adult population.

In the USA, the prevalence of high cholesterol has increased from 50 million in 1990 to 65 million in 2000. Reported differences by gender and race are small. The increasing prevalence is primarily a consequence of trends for the population to become older and more obese of increasing survival of hypertensive patients as a result of improved lifestyles or more effective drug therapy.

It is very important to refer that the WHO MONICA Project (2004), sample mainly represents populations from developed countries. Data from national surveys in six European countries, performed in the 1990s, using similar sampling and reporting techniques, estimated the prevalence of high blood pressure/hypertension as 38% in Italy, 38% in Sweden, 42% in England, 47% in Spain, 49% in Finland and 55% in Germany. In Portugal, data suggest that 3,311,830 people have high blood pressure/hypertension (42.1%). As a result of progressive urbanization and westernization of their lifestyle, developing countries are now undergoing an epidemiological transition. These changes are leading to a new epidemiological situation with a decline in infectious diseases and emergence of diabetes and cardiovascular diseases.

However, the reported hypertension prevalence was 27.2 in India, 40.6% in Syria 23.9% among men and 13.7% among women in Vietnam and 27.1% among men and 30.2% among women in Tanzania. These values are lower compared with high blood pressure/hypertension prevalence in developed world, but the global tendency is for these values to increase. Differences in hypertension prevalence are not only present between countries, but also between racial or ethnic groups. The prevalence among U.S. Blacks is higher than in Whites and Mexican-Americans in both genders and all ages.

Smoking

Cigarette smoking was first pointed out as a risk factor for diabetes in men in the late 1980s and later confirmed in large cohort studies by Monica Augsburg (2001), in both men and women and most prominently amongst heavy smokers. A recent paper summarizes the 15 published prospective epidemiologic studies on this subject.

Recently, a large Swedish cross-sectional study showed that the prevalence of type 2 diabetes, diagnosed by an oral glucose tolerance test (OGTT), was equally increased for smokers and snus users with high tobacco consumption compared with non tobacco users.

Finally, l believe from the studies above that the biggest change will occur in the developing world and developing countries if investment is made in both education on diabetes and o the elderly.

METHODOLOGY AND DATA ANALYSIS

The study would have considered all the main hospitals in the country but Ladoke Akintola Teaching Hospital was sampled out for the research.

The data used for the study was secondary data. This was taken from the physiotherapy department of the Ladoke Akintola Teaching Hospital.

The researcher visited the physiotherapy department of the Ladoke Akintola Teaching Hospital in Ogbomoso to take the following information from the folders of patients with Cerebral Vascular Accident (CVA) and diabetes. Cholesterol Level, glucose level, smoking status,family history, hypertension and diabetes status.

The software package used for the data analysis was SPSS version 21.

Categorical variable

These are variables that places individuals literally into categories and cannot be quantified in a meaningful way. Example would be diseases like diabetes, occupation, cholesterol level, gender etc.

Analysis of categorical data involves the use of data tables. A *two-way* table present categorical data by counting the number of observations that fall into each group for two variables one divided into rows and the other into columns.

LOGISTIC REGRESSION

Introduction

Regression methods have become an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables. It is often the case that the outcome variable is discrete, taking on two or more possible values. Over the last decade the logistic regression model has become, in many fields, the standard method of analysis in this situation. Hosmer and Lemeshow, (2000).

What distinguishes a logistic regression model from the linear regression model is that the outcome variable in logistic regression is *binary or dichotomous*. This difference between logistic and linear regression is reflected both in the choice of a parametric model and it assumptions. Once this difference is accounted for, the methods employed in an analysis using logistic regression follow the same general principles used in linear regression.

GENERALIZED LINEAR MODELS (GLMS) AND LOGISTIC REGRESSION.

The logistic regression model is an example of a broad class of models known as Generalized Linear Models (GLM). For example, GLMs also include linear regression, ANOVA, Poisson regression, etc. There are three components to a GLM:

Random Component – refers to the probability distribution of the response variable (Y); e.g. binomial distribution for *Y* in the binary logistic regression.

Systematic Component - refers to the explanatory variables ($x1$, $x2$,...,xk) as a combination of linear predictors; e.g. β0 + β1x1 + β2x2 as we have seen in logistic regression.

Link Function, *η or g(μ)* - specifies the link between random and systematic components. It says how the expected value of the response relates to the linear predictor of explanatory variables;

e.g. $\eta = logit(\pi)$ for logistic regression.

BINARY LOGISTIC REGRESSION

Models how binary response variable depends on a set of explanatory variable

Random component: The distribution of Y is *Binomial*

Systematic component: *X*s are explanatory variables (can be continuous, discrete, or both) and are linear in the parameters β0 + βxi + ... + β0 + βxk

Link function: *Logit*

η = Logit (□)= Log(□ 1− □)

The importance of this transformation is that g(x) has many of the desirable properties of a linear regression model. The logit, g(x), is linear in its parameters, may be continuous, and may range from –∞ to +∞, depending on the range of □

Fitting the Single Logistic Regression Model

The method of estimation used in fitting the logistic regression model is the maximum likelihood.

In order to apply this method we must first construct a function, called the *likelihood function.* This function expresses the probability of the observed data as a function of the unknown parameters. The *maximum likelihood estimators* of these parameters are chosen to be those values that maximize this function. Thus, the resulting estimators are those which agree most closely with the observed data. We now describe how to find these values from the logistic regression model.

If Y is coded as 0 or 1 then the expression □ (x) given in equation (3.0) provides (for an arbitrary value of □ = (β0, β1), the vector of parameters) the conditional probability that Y is equal to 1 given x. This will be denoted as P(Y = 1│x). It follows that the quantity 1 - □ *(x)* gives the conditional probability that Y is equal to zero given x, P(Y = 0│x). Thus, for those pairs (□□,), where □□=1, the contribution to the likelihood function is (□□,) and for those pairs where □□=0, the contribution to the likelihood function is 1-(□□,), where the quantity □(□□) denotes the value of □ (x) computed at □□.

A way to express the contribution to the likelihood function for the pair (□□,) is through the expression

(□□,)[1− (□□,)]1−□□(3.8)

Since the observations are assumed to be independent, the likelihood function is obtained as the product of the terms given in equation (3.8) as

l(□) = □(□□,)□□[1− □(□□,)]1−□□□□=1 (3.9)

The principle of maximum likelihood states that we use as our estimate of □ the value which maximizes the expression in equation (3.8) However, we will work with the log of equation (3.9). This expression, the *log likelihood* is given as:

L(□) = In[*l(□)*] = □□In[□(□□,)]□□=1 + (1-□□)In[1- □(□□,)] (3.10)

To find the value of □ that maximizes L(□) we differentiate L(□) with respect to β0,β1 partially and set the resulting expressions equal to zero.

These equations, known as the *likelihood equations,* are;

[□□−□(□□)]=0 (3.11)

and

xi[□□−□(□□)]=0 (3.12)

.

As in the univariate model, the solution of the likelihood equation requires special statistical software packages. In calculating the standard error, we will have to find the estimates of the variance and covariance of our coefficients. The method of estimating variances and covariance of the estimated coefficients follows from the theory of maximum likelihood estimation which states that the estimators are obtained from the matrix of second partial derivatives of the log likelihood function.

$$Y_i = \begin{cases} 1 \text{ diabetic} \\ 0 \text{ non diabetic} \end{cases}$$

Diabetic Status (Diabetic = 1; Non Diabetic = 0)

Glucose Level (High = 1; Low = 0)

Smoking (Yes = 1; No = 0)

Cholesterol level (High = 1; Low = 0)

Hypertensive (Yes = 1; No = 0)

Family history (Yes = 1; No = 0)

The dataset used for this paper was gotten from Ladoke Akintola University of Technology, Ogbomoso, Oyo state. The analysis was done using SPSS 21 package. This chapter uses the methodology of binary logistic regression analysis. The total number of subjects used was 200 patients. The response variable is the diabetic status.

GRAPHICAL DESCRIPTION OF DIABETIC STATUS ACROSS THE INDEPENDENT VARIABLES

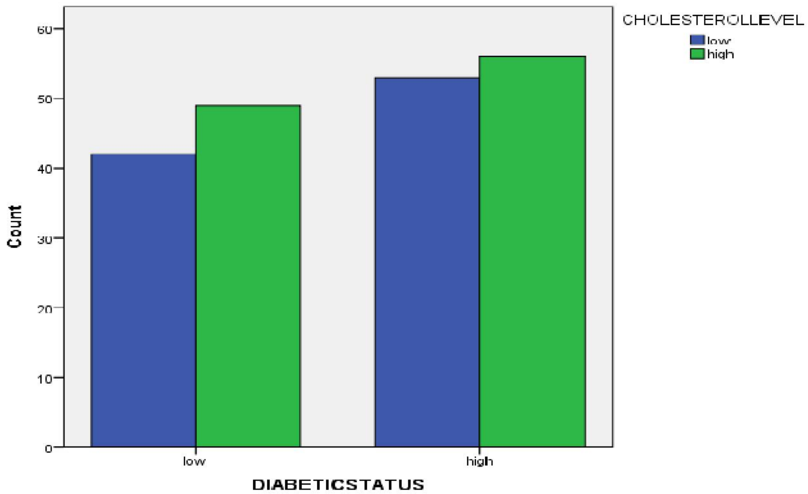
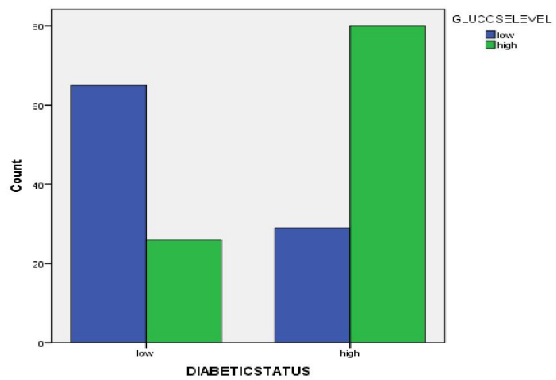
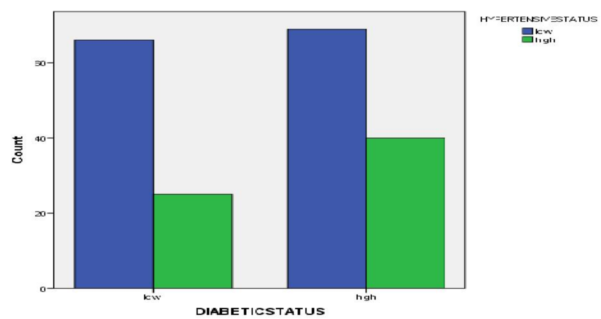


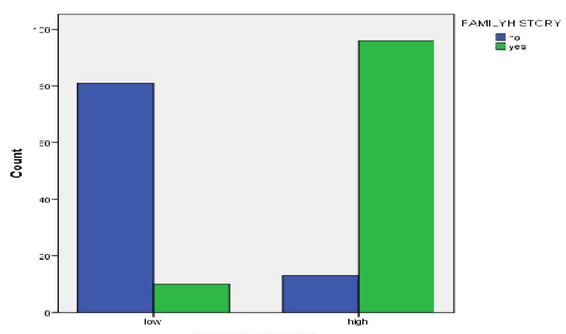
Figure 1: Multiple Bar Chart of Diabetic status across the Cholesterol level



:ose level



ertensive status



ily history

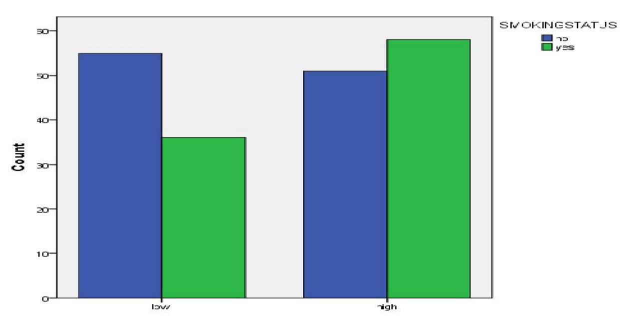


Figure 5: Multiple Bar Chart of Diabetic status across the Smoking status

The logistic regression model is:

The model was formulated as follows. For the literacy, let Y_i be the binary outcome diabetes, (diabetic/non-diabetic) for the individual I, $Y_i \sim$ Bernoulli (π_i).

$$\ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 Highcholesterollevel + \beta_2 HighGlucoseLevel + \beta_3 Highhypertensivestatus + \beta_4 Smoking + \beta_5 familyhistory + \varepsilon$$

The Full Model Assessment

Table 1: likelihood Ratio Test

Test	Chi-Square	Degree of freedom	P-value
Likelihood Ratio	154.680	5	0.000

The Table 1 displays the likelihood ratio test.
The likelihood ratio test is denoted with (G)

$$2log = \left(\frac{likelihood\ of\ the\ constant\ in\ the\ model}{likelihood\ of\ the\ overall\ model}\right) \sim x^2{}_{k,\alpha}$$

Where k is the number of predictors in the logistic regression equation.
Level of Significance (α) = 0.05
Decision Rule: Reject H_0 if G>$x^2{}_{k,\alpha}$, otherwise do not reject H_0 .
The tabulated value is $x^2{}_{k,\alpha} = x^2{}_{5,0.05=11.07\ 0}$
The hypothesis is:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \cdots = \beta_5 = 0$$
$$H_1: \beta_i \neq 0, i = 1,2,3,\ldots,5\ for\ at-leastonei$$

Decision Rule: Reject H_0 if G >$x^2{}_{k,\ \alpha}$, otherwise do not reject H_0
Interpretation: At-least one of the categories of the factors considered are significantly different from zero since the G >$x^2{}_{k,\ \alpha}$ I.e (154.680> 11.070) and thus, a binary logistic regression analysis can be carried out.
The maximum likelihood method is used to estimate the parameter estimates and the estimates are used to fit a binary logistic regression model for the dataset. Table 2 shows the output.

Parameter	Estimates (β)	Standard error	Wald value	D.f
Constant	-3.124	0.611	26.155	1
High Cholesterol level	-0.996	0.592	2.834	1
High Glucose level	1.793	0.497	13.025	1
High Hypertensive status	1.417	0.615	5.313	1
Family History	4.036	0.509	62.944	1
Smoking	1.085	0.501	4.698	1

TESTING FOR THE SIGNIFICANCE OF THE INDIVIDUAL REGRESSION PARAMETERS

Statement of Hypothesis:

$$H_0: \beta_i = 0 (for\ the\ individual\ parameter\ \beta_i) vs\ H_1: \beta_i \neq 0 (for\ the\ individual\ parameter\ \beta_i)$$

The test statistic for carrying out the test of significance of individual parameter is the Wald statistic, and it is given as:

$$W = \frac{\beta_i^2}{SE(\beta_i)^2} \sim x^2{}_{1,\alpha}$$

Level of Significance (α) = 0.05
Decision Rule: $Reject\ H_0\ if\ W > x^2{}_{1,\alpha}, otherwise\ do\ not\ reject\ H_0$
The tabulated value is ; = $x^2{}_{1,\alpha} = x^2{}_{1,0.05} = 3.841$
Interpretation: All the categories of the factors considered except high cholesterol level were significant in explaining the variation on diabetic status because $W > x^2{}_{1,\ \alpha}$ for all the factors.
The fitted Binary Logistic Regression Model With the significant predictors only

$$In\left(\frac{Diabeticpatients}{Non-diabeticpatients}\right) = -3.124 + 1.793 HighGlucoselevel + 1.417 HighHypertensivestatus + 1.085 Smoking + 4.036 familyhistory + \varepsilon$$

SUMMARY AND CONCLUSION

The data used for this paper consists of five factors namely cholesterol level, glucose level, hypertensive status, smoking status and family history. The paper consists of 400 patients whose information on the factors considered was complete.

The graphical representation of the diabetic status of patient across each of the factors was depicted to illustrate the frequency of each of the factors across the diabetic status.

All the independent variables except high cholesterol level were significant on predicting the outcome of a diabetic patient. The likelihood ratio test reveals that at least one of the predictor variables are significantly different from zero.

CONCLUSION

Based on the principal aim of fitting a binary logistic regression model for the factors associated with the diagnosis of diabetes. A fitted binary logistic regression model with the significant predictors only was given to estimate the odds of a person suffering or not suffering from diabetes. The model fitted shows that getting diabetes does not depend on the cholesterol level.

REFERENCES

[1]. Addo, J., Smeeth, L., &Leon, D.L.(2007). Hypertension in Sub Saharan Africa a systematic review. Hypertension. 50(6),1012-18.
[2]. Addo, J., Amoah, A.G., & Koram, K.A (2006). The changing pattern of hypertension in Ghana.A study of four rural communities in the GA district.Ethn Dis. 16(4), 89-99.
[3]. Agyeman, C. (2006). Rural and urban differences in Blood pressure and Hypertension in Ghana, West Africa.The Journal of Public Health. 120(6), 525-33
[4]. Agreti, A. (1990). Categorical Data Analysis, Wiley, Inc., New York
[5]. Ajani, U.A., Gaziano, J.M., Lotufo, P.A., Liu, S., Hennekens, C.H., Buring, J.E., et al. (2000). Alcohol consumption and risk of coronary heart disease by diabetes status.Circulation.102(5), 489-90.
[6]. Ali H. Mokdad, PhD; Earl S. Ford, MD, MPH; Barbara A. Bowman, PhD; William H. Dietz, MD, PhD; Frank Vinicor, MD, MPH; Virginia S. Bales, MPH; James S. Marks, MD, MPH, Prevalence of Obesity, Diabetes, and Obesity-Related Health Risk Factors, 2001
[7]. Anderson, T.W. (1984). An Introduction to Multivariate Statistical Analysis, Second Edition, Wiley, Inc., New York.
[8]. Bachand, A.M., and Hosmer, D.W. (1999). Defining confounding when the outcome is dichotomous, Submitted for Publication.
[9]. Baltazar, J.C., Ancheta, C.A., Aban, I.B., Fernando, R.E., &Baquila, M..M. (2004). Prevalence and correlation of diabetes mellitus and impaired glucose tolerance among adults in Luzon, Phillipines.Diabetes Research and Clinical Medicine.64(2), 107-552.
[10]. Basnet, S., & Sharma, S. (2000).Ageing Problem Two Sides of Coin.Kathmandu: Nepal Participatory Action Network.
[11]. Bedrick, E.J., and Hill, J.R. (1996). Assessing the fit of logistic regression models to individuals matched sets of case-control data. Biometrics,52, 1-9.
[12]. Beaty T. H. (Johns Hopkins U. School of Hygiene and Public Health, Baltimore, MD 21205),
[13]. J. V. Neel and S. S. Fajans. Identifying risk factors for diabetes in first degree relatives of non-Insulin dependent diabetic patients. Am J Epidemiol1982;115:380-97.
[14]. Bendel, R. B., and Afifi, A. A. (1977). Comparism of stopping rules in forward regression. Journal of the American Statistical Association,72, 46-53.
[15]. Burchfiel, C.M., Curb, J.D., Rodriguez, B.L., Yano, K., Hwang, L.J., Fong, K.O., et al. (1995). Incidence and predictors of diabetes in Japanese-American men.
[16]. The Honolulu Heart Program.An Epidermiology. 5(1), 33 - 43.
[17]. Campbell, N. R., Ashley, M. J., Carruthers, S.G., Lacourciere, Y., &Mckay, D .W. (1999).
[18]. Lifestyle modification to prevent and control hypertension. 3 recommendations on alcohol consumption. Canadian hypertension society, Canadian coalitation for high blood pressure prevention and control, laboratory center for Disease control and health, Canada, heart and stroke foundation of Canada. Canadian Medical Association Journal. 160(9) , 513-520.
[19]. Carballo, M. &Friedrik, S. (2006). Migration and Diabetes: the emerging challenge. Diabetes Voice. 51(2), 31-33.
[20]. Chao, W-H., Palta, M., and Young, T. (1997). Effect of omitted confounders in the

analysis of correlated binary data. Biometrics, 53, 678-689
[21]. Collett, D. (1991) Modelling Binary Data. Chapman & Hall, London.
[22]. Cox, D. R. (1970).The Analysis of Binary Data. Methuen, London.
[23]. Cox, D. R., and Snell, E. J. (1989).Analysis of Binary Data, Second Edition.Chapman& Hall, London.
[24]. Djousse, L., Biggs, M.L., Mukamal, K.J., & Sisicovic, D.S. (2007). Alcohol consumption and type 2 diabetes among older adults; The cardiovascular health study. Obesity 15(7), 1758-1765.
[25]. Dobson, A. (1990). An Introduction to Generalised Linear Models.Chapman& Hall, London.
[26]. Hosmer, D. W., Jovanovic, B., and Lemeshow, S. (1989). Best subsets logistic regression. Biometrics, 45, 1265-1270.
[27]. Hosmer, D.W., and Lemershow, S. (1980). A goodness-of-fit test for the multiple logistic regression model. Communications in Statistics, A10, 1043-1069.
[28]. Hosmer, T., Hosmer, D. W., and Klar .J. (1988). Goodness-of-fit testing for mulitiple logistic regression analysis when the estimated probabilities are small. Biometrical Journal, 30, 911-924.
[29]. Hosmer, D. W., Hosmer, T., Le Cessie, S., and Lemeshow, S. (1997).Acomparism of goodness-of-fit tests for the logistic regression model, Statistic in Medicine,8, 795-802.
[30]. Hosmer, D, W., and Lemeshow, S.(1999). Applied Survival Analysis: Regression Modelling of Time to Event Data.Wiley, Inc., New York.
[31]. Keil U, Liese AD, Hense HW, et al. Classical risk factors and their impact on incident non-fatal and fatal myocardial infarction and all-cause mortality in southern Germany. Results from the MONICA Augsburg cohort study 1984–1992. Eur Heart J 1998; 19: 1197–207.
[32]. Kim, S.M., Lee, J.S., Lee, J., Na, J.K., Han, J.H., & Yoon, D.K. (2006).Prevalence of diabetes and impaired fasting glucose in Korea. Diabetes Care.29(2), 226-32.
[33]. Lemeshow, S., and Hosmer, D .W.,(1982). The use of goodness-of-fit statistics in the development of logistic regression models. American Journal of Epidemiology, 115, 92-106.
[34]. Lemesho, S., and Hosmer, D. W., (1983).Estimation of odds ratios with categorical scaled covariates in multiple logistic regression analysis.American Journal of Epidemiology, 119, 147-151.
[35]. Lemeshow, S., Hosmer, D, W., Klar, J., and Lwanga, S.K. (1990).Adequacy of Sample Size in Health Studies.Wiley, Chichester.
[36]. Lemeshow, S., Teres, D., Avrunin, J. S., and Pastides, H. (1988).Predicting the outcome of intensive care uniypatients.Journal of the American Statistical Association, 83, 348-356.
[37]. McCullagh, P. (1986). The conditional distribution of goodness-of-fit statistics for discrete data.Journal of the American Statistical Assocition, 81, 104-107.
[38]. McCullagh, P., and Nelder, J. A. (1989). Generalized Linear Models, Second Edition. Chapman & Hall, London.
[39]. ang Y, Rimm EB, Stampfer MJ, Willett WC, Hu FB. Comparison of abdominal adiposity and overall obesity in predicting risk of type 2 diabetes among men. Am J Clin Nutr 2005; 81: 555–63.
[40]. World Health Organisation [WHO]. (1994). Prevention of diabetes mellitus: Report of a WHO study group. WHO technical report series 844; Geneva: Geneva.
[41]. Zozuliniska, D. &Wysoca, B.W. (2006). Type 2 diabetes mellitus as inflammatory disease. Diabete Research and Clinical Practice. 6(7).

10//24/2021