Faculty of Natural and Applied Sciences Journal of Mathematical and Statistical Computing
Print ISSN: 3026-8605
www.fnasjournals.com
Volume 1; Issue 2; March 2024; Page No. 37-42.

Study of the effects of Type I and Type II Diabetes using Logistic Regression Model

***Onu, O.H., Ejukwa, J.O., & Anyamele, G.A.**
Department of Mathematics/Statistics, Ignatius Ajuru University of Education, Port Harcourt, Nigeria

***Corresponding author email:** onu.obineke-henry@iaue.edu.ng

Abstract
The study of the effects of type I and type II diabetes using Logistic regression was considered. The data used was in the year 2021, obtained from the University of Port Harcourt Teaching Hospital (UPTH), Port Harcourt. The study employed, chi-square, p-value, goodness of fit test, and analysis of variance. It was observed that type II diabetes was more dangerous than type I. Also, the type 2 diabetes has a more negative effect on the patients than the type I. It was also found that both type I and type II are unsafe for one's health.

Keywords: Type I And Type II Diabetes, Binary, Probit, p-value.

Introduction

According to a study by the World Health Organisation (WHO,2016), the number of people diagnosed with diabetes is increasing globally. According to Nelson and Onu (2024), diabetes accounted for 3% of the approximately 36 million fatalities in 2008 that were reported to be caused by chronic non-communicable illnesses globally. In 2014, 422 million individuals, or around 8.5% of the global population, were living with diabetes. As a share of total mortality from chronic illnesses, diabetes has been steadily increasing in recent years. Elevated blood glucose levels are a hallmark of the metabolic disorder diabetes mellitus (DM). Numerous consequences, including high blood pressure, coronary heart disease, diabetes nephropathy, and diabetic foot, may develop from this chronic and lifelong condition. Current trends and statistics on diabetes indicate that the global prevalence of the disease will rise to 366 million by the year 2030. As many as 15.5% of Chinese adults may have diabetes at some point in their lives, with 60.7% of those people having never been diagnosed with prediabetes (Siyu et al., 2019)

Categories of Diabetes

Conditions caused by the illness determine the kind of diabetes that a person may have. The American Diabetes Association established a system for categorising the disease, which they call Type I, Type II, gestational diabetes, etc.

Type I diabetes

A chronic condition known as type I diabetes, which is dependent on insulin in the body, develops when the pancreas either produces insufficient insulin or no insulin at all. Insulin is the hormone that cells need to import sugar for energy synthesis. Type I diabetes may be caused by a number of things, including inherited traits and viral infections.
Type 1 diabetes is more common in children and teenagers, although it may also develop in adults.

Type I diabetes

The kind of diabetes that affects adults known as kind 2 diabetes does not rely on insulin, as stated by Parastoo and Ahmad (2016) and cited in Nelson and Onu (2024). It accounts for over 90% of all cases of diabetes and is among the most frequent forms. Similar to type 1 diabetes, type 2 diabetes occurs when the body either does not utilise its insulin effectively or does not create enough insulin from the pancreas. When insulin is either not produced or used by the body, glucose (sugar) builds up because it is unable to enter the cells. Doing this puts one's health at risk. A good diet, regular exercise, and general fitness may help one live with this illness for a long time, even if there is currently no treatment. You may need insulin therapy or medication if lifestyle changes alone are insufficient.

In his research on diabetes care in Damaturu, Nigeria, Umar (2018) used multivariate analytic procedures, including Fisher's technique. In a recent study, Onu et al. (2022) compared Poisson and Binary logistic regressions of type I and type II diabetes patients in Nigeria for dichotomous and non-dichotomous predictors. The study involved studying diabetic patients and examining binary logistic and Poisson regression models. Both type I and type II diabetes are harmful to patients, and the binary-logistic model was shown to be significant whereas the Poisson model was not. While these studies did look at the impact of type I and type II diabetes on Nigerians in 2021, they didn't use binary or Probit regressions. This painting was shown against this background. The purpose of this research is to study the effects of Type I and Type II Diabetes mellitus on Human beings using a logistics regression model.

To predict albuminuria in type II diabetics, Morteza et al. (2013) used a comparative technique and contrasted conditional logistic regression with neural networks.

To compensate for demographic and diagnostic effects, Sernyak et al. (2002) used logistic regression analysis to determine the chances ratio of a neuroleptic uncommon version and a diabetes diagnosis in each age group. The use of fuzzy neural networks has been enhanced for diabetes prediction (Thirugnanam et al., 2012).

Rather than measuring urine albumin, Marateb et al. (2014) proposed hybrid intelligence systems to identify microalbuminuria in type 2 diabetic patients.
Blood sugar regulation based on autonomous learning was suggested by Torkestani and Pisheh (2014) for type II diabetes.

Materials and Methods

Probit Regression

This is a member of the family of generalized linear models, like the logistic regression, its link function is given as

$$f(\mu_r) = \Phi^{-1}(p)$$

As mentioned in the equation, it makes use of an inverse normal link function. Since Y in a binary situation can only take on the values 0 and 1, we are interested in the relationship between a predictor and the likelihood that Y=1, but we can't utilise the probability as a function for reasons like these:

1. Probability can only be 0 or 1, that is to say, the right-hand-side of the equation can vary from $-\infty \ to \ \infty$
2. The relationship between probability and the predictors is not linear, rather, it is sigmoidal (S-shaped).

Consequently, a probability function is required to transform a probability into a value between -∞ and ∞, with a linear connection to the predictors. Consequently, logistic and Probit regression are ideal for these two purposes. Refer to Grace-Martin, Karen Karen

Logistic Regression Model

Onu et al. (2022) and Nelson and Onu (2024) both provide examples of logistic regression, a generalised linear model.

$$Logit(p) = \ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_{1i} + \cdots + \beta_k x_{ki}, i = 1, \dots, n \tag{1}$$

The link function of a logit regression is $f(\mu_r) = \ln\left(\frac{p}{1-p}\right)$

Where

$p = p_r(Y_i = 1)$ and

$$p = p_r\left(Y_i = \frac{1}{X}\right) = \frac{e^{\mu+\beta_1 x_{1i}+\cdots+\beta_k x_{ki}}}{1+e^{\mu+\beta_1 x_{1i}+\cdots+\beta_k x_{ki}}} \quad (2)$$

Logistic regression is quite similar to linear regression; however, the two methods vary in the coefficients that are used. Optimal event probability is maximised by logistic regression, but the sum of square errors is minimised by linear regression. For logistic regression, we utilise chi-square and Wald, but for linear regression, we use F and t-statistics.

In most cases, it is provided as:

$$\log\left(\frac{p}{1-p}\right) = \mu + \sum \beta X \quad (3)$$

X is a vector of independent variables

β is the vector of estimated parameters

P=is the likely outcomes or events.

Results

Logistic Regression: gender versus type I, type 2

***** WARNING * When the data are in the Response/Frequency format, the Residuals versus fits plot is unavailable.

Deviance Table

Source	DF	Adj Dev	Adj Mean	Chi-Square	P-Value
Regression	2	6.776	3.388	6.78	0.034
type I	1	1S.616	1.616	1.62	0.204
type 2	1	1.932	1.932	1.93	0.165
Error	21	26.495	1.262		
Total	23	33.271			

Model Summary

Deviance R-Sq	Deviance R-Sq(adj)	AIC
20.37%	14.35%	32.50

Coefficients

Cite this article as:
Onu, O.H., Ejukwa, J.O., & Anyamele, G.A. (2024). Study of the effects of type I and type II diabetes using the logistic regression model. *FNAS Journal of Mathematical and Statistical Computing*, *1*(2), 37-42.

Term	Coef	SE Coef	VIF
Constant	2.68	1.34	
type 1	-0.0598	0.0561	1.12
type 2	-0.0716	0.0548	1.12

Regression Equation

P(1) = exp(Y')/(1 + exp(Y'))

Y' = 2.68 - 0.0598 type 1 - 0.0716 type 2

Goodness-of-Fit Tests

Test	DF	Chi-Square	P-Value
Deviance	21	26.50	0.188
Pearson	21	23.82	0.302
Hosmer-Lemeshow	8	5.54	0.698

Discussion

With a p-value of 0.034, the logistic regression analysis demonstrated that the regression equation is statistically significant, indicating a meaningful relationship between the predictor variables and the response variable. The results highlighted that Type II diabetes is more severe and active compared to Type I diabetes, as evidenced by the adjusted mean square. The model explained 20.37 percent of the variance in the response variable (Gender), suggesting a moderate explanatory power.

The analysis confirmed that there is no presence of multicollinearity in the data, as indicated by the Variance Inflation Factor (VIF) and an Akaike Information Criterion (AIC) value of 32.50, both of which suggest a well-fitting model. Despite the negative impacts observed for all types of diabetes on the genders studied, the regression model particularly underscored that Type II diabetes has a more detrimental effect on individuals compared to Type I diabetes.

These findings are consistent with the conclusions drawn by Nelson and Onu (2024), reinforcing the assertion that Type II diabetes poses a greater risk and has more severe implications for affected individuals than Type I diabetes. The study adds to the growing body of evidence on the differential impacts of diabetes types on health outcomes across genders.

Conclusion

The study concludes that the logistic regression model is an effective and appropriate framework for investigating the impact of diabetes on human health. The analysis demonstrates that people of both sexes suffer from the detrimental effects of both Type I and Type II diabetes. However, the findings indicate that Type II diabetes is significantly more severe, posing greater health risks and leading to more serious complications than Type I diabetes. The model explained 20.37 percent of the variance in gender-related outcomes, highlighting its moderate explanatory power. The absence of multicollinearity in the data, supported by favourable VIF and AIC values, confirms the robustness of the model. These results underscore the critical need for targeted interventions and management strategies specifically tailored for those with Type II diabetes. This study highlights the value of using robust statistical models like logistic regression to uncover significant health disparities, guide effective public health policies, and develop better treatment plans and preventive measures. In addressing the specific needs of patients with Type II diabetes, healthcare professionals can improve the quality of life and health outcomes for this vulnerable population.

Recommendations

The study recommends the following to statisticians medical researchers and the general public that

1. Health workers should take the treatment of type I diabetes seriously to avoid going into the danger zone of type II.
2. The patients should adhere strictly to the advice of the health workers to avoid the severity of the illness.
3. The general public should ensure regular medical checkups to ascertain their health status on time, for this will solve the problem of diabetes if appropriate measures are taken.

References

Marateb, H. R., Mansourian, M., Faghihimani, E., Amini, M., & Farina, D. (2014). A hybrid intelligent system for diagnosing microalbuminuria in type 2 diabetes patients without having to measure urinary albumin. *Computers in biology and medicine*, *45*, 34-42.

Morteza, M. N. (2013). Inconsistency in albuminuria predictors in type II diabetes: a comparison between neural network and conditional logistic regression. *Translational Research*, *161*(5), 397-405.

Nelson, M., & Onu, O. H. (2024). Comparative study of the effects of Type I and Type II Diabetes on Human Health, using Statistical Models. *Indian Journal of Modern Research and Reviews*, *2(3)*, 1-5.

Onu, O. H., Amakuro, O. A., & Alabge, S. A. (2022). Study of Binary Logistic and Poisson Regression Models of Diabetic Patients in Nigeria using Dichotomous and Non- Dichotomous Predictors. *Asian Journal of Probability and Statistics, 17*(3): 37-48.

Sernyak, M. J., Leslie, D. L., Alarcon, R. D., Losonczy, M. F., & Rosenheck, R. (2002). Association of diabetes mellitus with use of atypical neuroleptics in the treatment of schizophrenia. *American Journal of Psychiatry*, *159*(4), 561-566.

Siyu L., Yue, G., Yuhang S. L, Min, J. L., & Pinghui, S. (2019); Application of three statistical models for predicting the risk of diabetes. *BCM Endocrine Disorders, 126.*

Thirugnanam, M., Kumar, P., Srivatsan, S. V., & Neriesh, C. R. (2012). Improving the prediction rate of diabetes diagnosis using fuzzy, neural network, case based (FNC) *Approach. Procedia Engineering*, *38*, 1709-118.

Torkestani, J. A., & Pisheh, E. G. (2014). A learning automata-based blood glucose regulation mechanism in type 2 diabetes. *Control Engineering Practice*, *26*, 151-159.

Umar, Y. M., Abdulrahman, M., & Maigana, A. B. (2018). Statistical Analysis on Tackling Diabetic Patients among Adults in Damaturu, Nigeria. International Journal of Innovative Research & Development, 7(10), 128-132

University of Port Harcourt Teaching Hospital (2021). Data of type I and type II diabetes from 2020 to 2021

World Health Organization (2016). Global report on diabetes.

Data of type I and type II diabetes for 2021

Data for 2020

Data for 2021

type I	type 2	Gender
20	16	1
10	3	1
12	9	1
10	7	1
4	1	1
13	4	1
15	3	1

12	10	1
21	10	1
9	3	1
15	3	1
12	7	1
33	14	0
85	24	0
30	2	0
17	8	0
10	1	0
51	12	0
7	12	0
20	3	0
25	12	0
25	5	0
20	8	0
26	2	0

Source: UPTH Port Harcourt.