

Empirical Study of Socio-Demographic Characteristics of Diabetes Mellitus Patients Using Cochran-Q-Test

Olanrewaju S.O. & Yahaya H.U.

Department of Statistics, University of Abuja, Abuja.

Email: olanrewaju.samuel@uniabuja.edu.ng

Abstract: This study examined the relationship between socio-demographic characteristics of patients suffering with diabetes mellitus disease in Gwagwalada Area Council of Abuja using Cochran-Q test. The data was gotten from the Records Department, University of Abuja Teaching Hospital, Gwagwalada, which ranges from August 2006 to December 2016. Cochran-Q statistic is a chi-square variation formed by a ratio of the variation across test variables to the variation within test variables. The degree of freedom is approximately chi-square, which is equal to the number of test variables minus 1. The results show that there are significant relationships between age, gender, complication caused by diabetes and mortality rate at 5% level of significance. A further test of pair-wise comparison analysis of the socio-demographic characteristics was carried out with Bonferroni correction which shows that there is a relationship between age and mortality status, between age and complication caused by diabetes mellitus disease and between complication caused by diabetes mellitus disease and mortality status, but no relationship between gender and mortality status nor between gender and complication caused by diabetes mellitus disease.

Keywords: Diabetes mellitus disease, Empirical Study

1. INTRODUCTION

14th of November each year is the world diabetes mellitus day, a day set aside for the awareness and prevention of the incurable disease. The term “diabetes” or “to pass through” was first used by the Greek Apollonius of Memphis. Diabetes was one of the first diseases described with Egyptian manuscript from 1500BCE mentioning “too great emptying of urine“. Indian physician around the same time identified the disease and classified it as madhumeha or “honey urine” noting that the urine would attract ants. The term “mellitus” or “from honey” was added by the Briton John Rolle in the late 1700’s to separate the condition from diabetes insipidus, which is also associated with frequent urination. While many measures were tried, effective treatment was not developed until the early part of 20th century, when Canadian Frederick Banting and Charles Herbert Best developed insulin in 1921 and 1922. This was followed by the development of insulin NPH in the 1940’s. Type 1 and type 2 diabetes mellitus were identified as separate conditions for the first time by the Indian physicians Sushruta and Charaka in 400-500CE with type 1 associated with youth and type 2 with being overweight. Diabetes mellitus disease is a disease assuming an epidemic proportion worldwide. Presently about 240 million people worldwide are said to be suffering from the disease and this figure is projected to get to about 340 million by the year 2030. Overall it is estimated that 8%-10% of people over 50years worldwide have diabetes and 40% will die from kidney disease and 60% from cardiovascular causes (Unadike, 2011). In Nigeria, the national prevalence of diabetes mellitus is put at 2.2% and this continues to be on the increase. Odeniye (2012), a consultant endocrinologist of the Faculty of Clinical Sciences, Collage of Medicine, Lagos University Teaching Hospital (LUTH) disclose at a health talk entitled “Diabetes Mellitus: An Emerging Threat to Human Race.” said that between 7 and 10 million Nigerians are suffering from diabetes “three out of four people who live with diabetes mellitus live in low-income and developing countries and about 98per cent of Africans would be diabetic in the next twenty years.” “In Nigeria, a life is lost every eight second from diabetes, while a leg is amputated every minute,” he said.

Whenever numbers are collected and compiled regardless of what they represent, they become statistics. This information is considered as statistics but in the real sense statistics has so many meanings from inception. The word ‘statistics’ comes from the Italian word statista (meaning “statesman”) or the German word ‘statistik’ which means a political state. It was first used by Professor Gottfried Achenwall (1719-1772), a professor in Marlborough in 1749 to refer to the subject matter as a whole. Achenwall defined statistics as “the political science of the several countries”. The word statistics appeared for the first time in the famous book, Elements of Universal Erudition by Baran J. F. Von Bielfeld, translated by W. Hooper M. D. (3 vols, London, 1770). One of its chapter is entitled ‘statistics’ and contains a definition of the subject as “the science that teaches us what is the political arrangement of the modern states of the known world”. There have been many definitions of the term ‘Statistics’ indeed scholarly articles have carefully collected together hundreds of definitions. Some have defined statistics as statistical data (plural sense) where as others as statistical methods (singular sense). It is probably more common to refer to data in qualitative form as statistical data. But not all numerical data is statistical and hence it is necessary to examine a few definitions of statistics to understand the characteristics of statistical data. Yule and Kendall defined statistics thus: “By statistics we mean quantitative data affected to a marked extent by multiplicity of causes “. This is definition less comprehensive than the one given by Prof. Horace Secrist who defined statistics as

follows: “By statistics we mean aggregates of facts affected to a marked extent by multiplicity of causes, numerical expressed, enumerated or estimated according to reasonable standard of accuracy, collected in a systematic manner for a predetermined purpose and placed in relation to each other.” The large volume of numerical information gives rise to the systematic methods which can be used to organize, present, analyse and interpret the information effectively. Statistical methods are primarily developed to meet this need. But what are statistical methods? The term statistics in this sense too has been defined differently by different writers. A few definitions are as follows. Prof. A. L. Bowley has given some definitions. At one place he says, “Statistics may be called the science of counting” this definition is too narrowed because it covers only one aspect of science namely the collection of data. Other aspects like analysis, presentation, interpretation etc are completely ignored. At another place Bowley says “Statistics may rightly call the science of average.” This definition still is not satisfactory because average is only one of the devices used in statistical analysis. The other devices like depression, skewness, correlation, etc. are not at all covered by the definition. Still another definition by the same author is “statistics is the science of the measurement of social organism, regarded as a whole manifestation.” This definition again is inadequate because it confines the application of statistics only to sociology, i.e man and his activities. Bowley himself realise this when he remarked “Statistics cannot be confined to any one science.” According to Berenson and Levin “The science of statistics can be viewed as the application of the scientific method in the analysis of numerical data for purpose of making rational decisions.

Croxton and Cowden have given a very simple and concise definition of statistics. In their view “Statistics may be define as the collection, presentation, analysis and interpretation of numeric data.” However to the above definition one more stage may be added and that is organization of data. Thus statistics may be defined as the science of collection, organization, presentation, analysis and interpretation of numeric data.

The researcher will use Cochran Q test a tool under nonparametric test in statistics to check the relationship between the socio-demographic data of in-patients suffering with diabetes mellitus disease with the assistance of SPSS software.

2. OBJECTIVE/AIM OF THE STUDY.

The researcher will discuss the application and importance of Cochran Q test under non parametric test, to test hypothesis in health related data.

1. To determine if there is relationship between age, gender, mortality status and complication caused by diabetes mellitus.
2. To check the relationship between age, complications caused by diabetes mellitus disease and mortality status as it relate with diabetes mellitus disease.
3. To check the relationship between gender, complications caused by diabetes and mortality status.

3. RESEARCH METHODOLOGY

The method of data analysis employed the non-parametric statistics. The package use is the Statistical Package for Social Science and the tool is Cochran Q test. Cochran Q test assumes that there are $c \geq 2$ experimental treatments and that the observations are arrange in blocks; The sampling technique used for this research work is the stratified sampling technique in which the diabetic patient medical data was classified into types then after that simple random sampling was used to select appropriate sample sizes independently from each stratum or subgroup. The size of the population was measure through the use of years from august 2006 to august 2016.

HYPOTHESIS

H_0 : The treatments are equally effective,

H_a : There is a difference in effectiveness among treatment

Let $X_{ij} = 1$ is effective (present) and 0 otherwise. Let $p_{ij} = P(X_{ij}=1)$ so that $X_{ij} \sim \text{binomial}(1, p_{ij})$. Under H_0 , $p_{i1}=p_{i2}=\dots=p_{ik}$ (say) for $i=1,2,\dots,r$, for r is the number of blocks. Under H_0 , each $X_{ij} \sim \text{binomial}(1, p_i)$ and upon applying Liapounov's central limit theorem, when r is large, the column totals have an approximate normal distribution: $C_j = \sum_{i=1}^r X_{ij} \sim \text{Normal}$. We then standardize the C_j , square the standardized result and add the squares to obtain a test statistics that has an approximate χ^2 distribution.

In standardizing the C_j , we use the following estimates of $E(C_j)$ and $\text{Var}(C_j)$. The mean $E(C_j)$ is estimated by the sample mean $\frac{\sum C_j}{c} = \frac{N}{c}$, where N denotes the total number of 1's. Since $\text{Var}(C_j) = \sum_{i=1}^r p_{ij}(1-p_{ij})$, we can estimate it by

$\sum \frac{R_r}{c} \left(1 - \frac{R_r}{c}\right) \left(\frac{c}{c-1}\right)$ where $\frac{R_r}{c}$ is the natural estimate of p_i under H_0 and where $\left(\frac{c}{c-1}\right)$ is a correction factor that safeguards against type one errors. In estimating the mean and variance in this context, we lose a degree of freedom for the χ^2 distribution Test

Statistic.
$$Q = \sum_{j=1}^c \left(\frac{c_j - \frac{N}{c}}{\sqrt{\frac{c}{c-1} \times \frac{R_i}{c} \left(1 - \frac{R_i}{c}\right)}} \right)^2 = c(c-1) \frac{\sum_{j=1}^c \left(c_j - \frac{N}{c}\right)^2}{\sum R_i(c-R_i)}$$

or equivalently
$$Q = \frac{c(c-1) \sum_{j=1}^c c_j^2 - (c-1)N^2}{cN - \sum_{i=1}^r R_i^2}$$

This, under H_0 has an approximate χ^2 distribution with $c-1$ degree of freedom. Where

c is the number of treatments

C_j the column total for the c^{th} treatment

r is the number of blocks

R_i Is the row total for the r^{th} block

N is the grand total

Assumptions

Cochran Q test is based on the following assumption:

- 1) A large sample approximation; in particular, it assume that r is large.
- 2) The blocks were randomly selected from the population of all possible blocks
- 3) The outcomes of the treatments can be coded as binary responses (i.e., a “0” or “1”) in a way that is common to all treatments within each block.

4. RESULT PRESENTATION AND ANALYSIS

The data that will be used for the research work is a secondary data which was gotten from the records unit of University of Abuja Teaching Hospital Gwagwalada. The data is on diabetic patient’s medical information based on age, gender, mortality status and complication caused by diabetes mellitus disease.

4.1 Frequency Table

The table summarize the number of observations of possible outcome 1 or 0 of each status.

- i) Age: 1 for Adult and 0 for Young.
- ii) Gender: 1 for Male and 0 for Female.
- iii) Complication: 1 for present and 0 for absent.
- iv) Mortality: 1 for improve and 0 for died.

4.2 Contingency table

Table 1.0.1
Gender * Complications caused by Diabetes Mellitus Disease
Crosstabulation

Count		Complications caused by Diabetes Mellitus Disease		Total
		ABSENT	PRESENT	
Gender	FEMALE	68	134	202
	MALE	99	193	292
	Total	167	327	494

Table 1.0.2

Gender * Mortality Status Crosstabulation

Count		Mortality Status		
		DIED	IMPROVE	Total
Gender	FEMALE	93	109	202
	MALE	134	158	292
	Total	227	267	494

Table 1.0.3

Age * Complications caused by Diabetes Mellitus Disease Crosstabulation

Count		Complications caused by Diabetes Mellitus Disease		Total
		ABSEND	PRESEND	
Age	YOUNG	1	11	12
	ADULT	166	316	482
	Total	167	327	494

Table 1.0.4 Age * Mortality Status Crosstabulation

Count		Mortality Status		Total
		DIED	IMPROVE	
Age	YOUNG	5	7	12
	ADULT	222	260	482
	Total	227	267	494

4.3 Analysis

4.3.1 Analysis between Age, Gender, Complication caused by diabetes mellitus disease and Mortality status

Table 2.0.0: Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Age	494	.9757	.15411	.00	1.00
Gender	494	.5911	.49213	.00	1.00
Complications caused by Diabetes Mellitus Disease	494	.6619	.47353	.00	1.00
Mortality Status	494	.5405	.49886	.00	1.00

Table2.0.1

Frequencies

	Value	
	0	1
Age	12	482
Gender	202	292
Complications caused by Diabetes Mellitus Disease	167	327
Mortality Status	227	267

Table 2.0.2 Test Statistics

N	494
Cochran's Q	2.499E2 ^a
Df	3
Asymp. Sig.	.000

a. 1 is treated as a success.

From table 2.0.2 we can be deduced that Age, gender, complications caused by diabetes mellitus and mortality status are related.

Table 3.0.0: Analysis between Age, Complication caused by diabetes mellitus disease and Mortality status. Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Age	494	.9757	.15411	.00	1.00
Complications caused by Diabetes Mellitus Disease	494	.6619	.47353	.00	1.00
Mortality Status	494	.5405	.49886	.00	1.00

Table 3.0.2 Test Statistics

N	494
Cochran's Q	2.231E2 ^a
Df	2
Asymp. Sig.	.000

a. 1 is treated as a success.

From table 3.0.2 Age, complication caused by diabetes mellitus and mortality status are related.

Table 4.0.0: Analysis between Mortality status, Complication caused by diabetes mellitus disease and Gender. Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Mortality Status	494	.5405	.49886	.00	1.00
Complications caused by Diabetes Mellitus Disease	494	.6619	.47353	.00	1.00

Table 4.0.0: Analysis between Mortality status, Complication caused by diabetes mellitus disease and Gender. Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Mortality Status	494	.5405	.49886	.00	1.00
Complications caused by Diabetes Mellitus Disease	494	.6619	.47353	.00	1.00
Gender	494	.5911	.49213	.00	1.00

Table4.0.2:Test statistics

N	494
Cochran's Q	14.891 ^a
Df	2
Asymp. Sig.	.001

a. 1 is treated as a success.

From table 4.0.2

Gender, complication caused by diabetes mellitus and mortality status are related.

4. 4 Pair Wise Comparison Using Bonfernni Correction

There was no enough statistical evidence based on this data to accept the null hypothesis, but concluded that there was relationship. However the overall significance does not provide specific information about the pattern of relationship. When completing multiple comparisons it is important to protect against Alpha inflation (increase chance of type 1 error as the number of comparisons increase). In this analysis since SPSS provide an exact p-value we can apply Bonfernnization of the p-value <rule. Our goal in applying the rule is to have no more than a 5% chance of a type 1 error across the pair wise comparison. The “Bonfernnized” p-value will be used to make each of our pair wise comparison. In order to obtain the pair wise comparison, the number of the pair wise comparison is computed as $\frac{c(c-1)}{2}$, which for this analysis is $\frac{3(3-1)}{2} = 3$. The Bonfernnized p-value would be $\frac{0.05}{3} = 0.0$

4.4.1 Analysis between Mortality status and Age

Table 5.0.0: Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Mortality Status	494	.5405	.49886	.00	1.00
Age	494	.9757	.15411	.00	1.00

Table5.0.2:Test statistics

N	494
Cochran's Q	2.019E2 ^a
Df	1
Asymp. Sig.	.000

a. 1 is treated as a success.

From table 5.0.2

Level of significance (Bonfernnization p-value) is 0.0167.

We have no reason to accept the null hypothesis.

Conclusion

Age and mortality status are related.

Table 6.0.0: Analysis between Complication caused by diabetes mellitus disease and Age Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Complications caused by Diabetes Mellitus Disease	494	.6619	.47353	.00	1.00
Age	494	.9757	.15411	.00	1.00

Table6.0.2:Test statistics

N	494
Cochran's Q	1.357E2 ^a
Df	1
Asymp. Sig.	.000

a. 1 is treated as a success.

From table 6.0.2

Level of significance (Bonfermnization p-value) is 0.0167.

Age and complications caused by diabetes are related.

4.4.3 Analysis between Mortality status and Gender

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Mortality Status	494	.5405	.49886	.00	1.00
Gender	494	.5911	.49213	.00	1.00

Table7.0.2 Test Statistics

N	494
Cochran's Q	2.572 ^a
Df	1
Asymp. Sig.	.109

a. 1 is treated as a success.

From table 7.0.2

Level of significance (Bonfermnization p-value) is 0.0167.

Gender and mortality status are not related

4.4.4 Analysis between Complication caused by diabetes mellitus disease and Gender

Table 8.0.0: Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Complications caused by Diabetes Mellitus Disease	494	.6619	.47353	.00	1.00
Gender	494	.5911	.49213	.00	1.00

Table8.0.2:Test statistics

N	494
Cochran's Q	5.258 ^a
Df	1
Asymp. Sig.	.022

a. 1 is treated as a success.

From table 8.0.2

Level of significance (Bonfermnization p-value) is 0.0167.

Gender and complication caused by diabetes mellitus disease are not related.

4.4.5 Analysis between Complication caused by mellitus disease and Mortality status

Table 9.0.0: Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Complications caused by Diabetes Mellitus Disease	494	.6619	.47353	.00	1.00
Mortality Status	494	.5405	.49886	.00	1.00

Table 9.0.2: Test statistics

N	494
Cochran's Q	14.062 ^a
Df	1
Asymp. Sig.	.000

a. 1 is treated as a success.

From table 9.0.2

Level of significance (Bonferroni p-value) is 0.0167.

Complications caused by diabetes mellitus and mortality status are related.

5. SUMMARY

The results of this project work on diabetes mellitus disease with their socio-demographic data will be reported statistically from a Cochran Q test format a non-parametric test. The results were obtained with the assistance of a computer program known as statistical package for social science (SPSS).

In the analysis the null hypothesis was rejected at 0.05 level of significance, which shows that based on these data we have sufficient statistical evidence to conclude that there is relationship in at least two socio-demographic data.

Prior to one of the advantages of Cochran Q test been a probe of its own, it was employed for the pair-wise comparison but in order to protect against alpha inflation, the p-value was bonferronized so as not to have more than a 5% chance of a type 1 error across the pair-wise comparison. There was relationship between age and mortality status, but no relationship between gender and mortality status also there was relationship between age and complication caused by diabetes mellitus, but no relationship between gender and complication caused by diabetes mellitus, whereas complication caused by diabetes mellitus and mortality status are related.

6. CONCLUSION AND RECOMMENDATION

Based on the results obtained from the project work, the following recommendations are made. In conclusion in order to prevent, delay or manage the chronic incurable disease among all age group there should be paramedic and mobile clinic or laboratory on the world awareness day of diabetes mellitus disease for people to be screen and individual at high risk of developing diabetes mellitus need to become aware of the benefit of modest weight loss and to participate in moderate physical activity. Adequate hospital and facility should be provided both in the cities and village and there should be free medical check-up in hospitals so that people can be screen at any time.

REFERENCES

[1] Bain, K. (2001). Depression and diabetes, Health and Hygiene 12(11). Black, J.M. and Matassarini-Jacobs, E. (1993). Lukmann and Sorensen's medical surgical nursing: a psycho-physiologic approach. 4th edition. Philadelphia: WB Saunders.

[2] Champion D. J. and Hertley R.D. (2008). Statistics for Criminal Justice and Criminology 3rd edition. Prentice Hall.

[3] Chinenye, S., & Ogbera, A. O. (2013). Socio-cultural aspects of diabetes mellitus in Nigeria. Journal of Social Health and Diabetes, 1(1), 15.

[4] Crowther, N. J. and Van der Merwe, M.T.(2001). Ethnicity and pathogenesis of obesity and type 2 diabetes in S.A CME 19(8):568-575.

[5] Gupta S. P. (2010). Statistical methods 39th edition Pp 3-9.

[6] Higgins J.P.T, Thompson S.G. (2002). Quantifying heterogeneity in a meta-analysis Statistics in Medicine

[7] Issa, B. A., & Baiyewu, O. (2006). Quality of Life of Patients with Diabetes Mellitus in a Nigerian Teaching Hospital. Hong Kong Journal of Psychiatry, 16(1).

[8] Mealey, B. L. and Ocampo G. L. (2007). Diabetes mellitus and periodontal disease. Periodontology 2000, 44: 127-153.

[9] Odeniye I. (2012). A Consultant endocrinologist of the faculty of clinical science, collage of medicine, Lagos University Teaching Hospital (LUTH). 'Diabetes Mellitus: An Emerging Threat to Human Race', captured in the tribune newspaper.

- [10] Ogbera, A. O., & Ekpebegh, C. (2014). Diabetes mellitus in Nigeria: The past, present and future. *World journal of diabetes*, 5(6), 905.
- [11] Phipps, W.J., Long, B.C. and Wood, N.F. (1987). *Medical-Surgical nursing: concept and clinical practice* 3rd edition. St Louis: Mosby.
- [12] Royle J. R. And Walsh M. (1992). *Watson's medical- surgical nursing and related physiology* 4th edition. London Bailliere Tindall.
- [13] Smeltzer, S.G. and Bare B.G. (1992). *Burner and Suddath's text book of medical nursing* 7th edition.
- [14] Smeltzer, S.G. and Bare B.G. (1996). *Burner and Suddath's textbook of medical nursing* 8 edition. Philadelphia.
- [15] Unadike B. (2011). University of Uyo Teaching Hospital Uyo, 'The burden of diabetes mellitus' captured in the Nation newspaper.
- [16] [En.wikipedia.org/./diabetes-mellitus](http://en.wikipedia.org/./diabetes-mellitus).
- [17] Frank.mtsedu././6604COCH.pdf
- [18] psych.unl.edu././hccohran
- [19] psych.unl.edu././kcochrans.pdf
- [20] www.vsni.co.uk././CochranQ.htm
- [21] www.staff.ncl.ac.uk././who_dcm.htm