Journal of the Nigerian Association of Mathematical Physics
Volume 51 (*May, 2019 Issue*), pp119 – 124

SURVIVAL ANALYSIS OF DIABETIC PATIENTS USING KAPLAN- MEIER PRODUCT LIMIT ESTIMATION FUNCTION: CASE STUDY

Nurudeen A. A., Kolawole I. A. and MUDI T.A.

Department of Mathematics and Statistics, Kaduna Polytechnic, Kaduna

Abstract

This study examines the survival times of diabetic patients attending two different diabetes centers. The data used were extracted from clinical studies-Usmanu Danfodiyo University Teaching Hospital, (UDUTH), Sokoto and Barau Dikko Teaching Hospital, (BDTH) Kaduna from 2012-2017, the population-based survey of known diabetes prevalence with a subsequent 6-year mortality follow-up. A cohort of 2,613 diabetic patients' folder was observed on 11th April, 2018 from two different sources: UDUTH, Sokoto and BDTH, Kaduna. As at the date stated, 1,193 patients in the entire cohort attended UDUTH while the remaining attended BDTH. The life status of the diabetic cohort was ascertained on 11th April, 2018. Using the Kaplan-Meier Product Limit Estimation Function, the mean survival time of attendees at UDUTH and BDTH is 59.913 and 59.812 months respectively. However, on the overall comparison, the survival was slightly higher at UDUTH, Sokoto diabetes center than at BDTH, Kaduna (P< 0.05).

***Keywords*:** survival time, diabetic, Cohort, Kaplan-Meier, attendance centre, clinical.

1.0 Introduction

Survival analysis techniques employ methods designed to investigate the amount of study time an experimental unit contributes to a study period from entry until event. The term "survival" may be misleading because the techniques are applicable to any well-defined event although traditionally death is the event of interest and the study period consisted of following the subject until death. Event in survival analysis (also referred to as endpoints or outcomes) are defined by a transition from one discrete state to another at an instantaneous moment in time. Example of events include months until onset of disease, days until remission after cancer therapy, years until stock market crash, hours until equipment failure, days until unemployment or time until failing or passing an examination [1].

Although the origin of survival analysis goes back to mortality tables from centuries ago, recent advancements in survival analytic techniques using non-parametric and semi-parametric approaches have allowed researchers flexibility in their work not properly seen within the confines of parametric methods. These methods have become popular over parametric methods due to the relatively robust modeling approaches without distributional assumptions on the survival times [2].

Diabetes is a chronic disease that arises when the pancreas does not produce enough insulin, or when the body cannot effectively use the insulin it produces. Insulin is a hormone made by the pancreas that enables cells to take in glucose from the blood and use it for energy. Failure to produce insulin, or of insulin to act properly, or both, leads to raise glucose (sugar) levels in the blood (hyperglycemia). This is associated with long term damage to the body and failure of various organs and tissues [3].

Diabetes comes from Greek, and it means a siphon. Aretus the Cappadocian, a Greek physician during the second century A.D., named the condition diabainein. He described patients who were passing too much water (polyuria) - like a siphon. The word became "diabetes" from the English adoption of the Medieval Latin diabetes. In 1675 Thomas Willis added mellitus to the term, although it is commonly referred to simply as diabetes. Mel in Latin means honey; the urine and blood of people with diabetes has excess glucose, and glucose is sweet like honey. Diabetes mellitus could literally mean "siphoning off sweet water". In ancient China people observed that ants would be attracted to some people's urine, because it was sweet. The term "Sweet Urine Disease" was coined [4].

Corresponding Author**:** Nurudeen A.A., Email: nurkarmeds@gmail.com, Tel: +2348032156015

There are three main types of diabetes, type I diabetes is sometimes called insulin-dependent, immune-mediated or juvenile-onset diabetes. It is caused by an auto-immune reaction where the body defence system attacks the insulin-producing cells. The reason why this occurs is not fully understood. People with type I diabetes produce very little or no insulin. The disease can affect people of any age, but usually occurs in children or young adults. People with this form of diabetes need injections of insulin everyday in order to control the level of glucose in their blood. If people with type I diabetes do not have access to insulin, they will die [5].

Type II diabetes account 90% of all cases of diabetes. Type II diabetes is sometimes called non-insulin dependent diabetes or adult-onset diabetes, and account for at least 90% of all cases of diabetes. It is characterized by insulin resistance and relative insulin deficiency, either of which may be present at the time that diabetes becomes clinically manifest. The diagnosis of type II diabetes usually occurs after the age of 40 but can be detected earlier, especially in population with high diabetes prevalence. Type II diabetes can remain undetected for many years and the diagnosis is often made from associated complications or incidentally through an abnormal blood or urine glucose test. It is often, but not always, associated with obesity, which itself can cause insulin resistance and lead elevated blood glucose levels [5].

Gestational diabetes (GDM) is a form of diabetes consisting of high blood glucose levels during pregnancy. It develops in one in a 25 pregnancies worldwide and is associated with complications in the period immediately before and after birth. GDM usually disappears after pregnancy but women with GDM they are at an increase risk of developing type II diabetes later in life. Approximately half of women with a history of GDM go on to develop type II diabetes within five to ten years after delivery [5].

Warning signs of diabetes, Individuals can experience different warning signs and sometimes there may be no obvious warning, but some of the signs of diabetes are commonly experienced:

Frequent urination, Excessive thirst, Increased hunger, Weight loss, Tiredness, Lack of interest and concentration, Vomiting and stomach pain (often mistaken as the flu), A tingling sensation or numbness in the hand or feet, Blurred vision, Frequent infections, Slow-healing wounds,. The onset of type I diabetes is usually sudden and dramatic while the symptoms can often be mild or absent in people with type II diabetes, making this type of diabetes gradual in onset and hard to detect. If you show these signs, consult a health professional

The risk factors, for type I diabetes are still being researched. However, having a family member with type I diabetes increases the risks for developing the condition, as do the presence of some genetic factors. Environmental factors increased height and weight development, increased maternal age at delivery, and exposure to some viral infections have also been linked to the risk of developing type I diabetes. Several risk factors have been associated with type II diabetes and include: Obesity, Diet and Physical inactivity, Increasing age, Insulin resistance, Family history of diabetes, Ethnicity. Changes in diet and physical activity related to rapid development and urbanization have led to sharp increases in the numbers of people developing diabetes. Pregnant women who are overweight, have been diagnosed with Impaired Glucose Tolerance (IGT), or have a family history of diabetes are all at risk of developing Gestational diabetes (GDM). In addition, having been previously diagnosed with gestational diabetes or being of certain ethnicities put women at risk of developing the condition [5]. The aim of the study is to analyze the survival times of diabetic patients at UDUTH and BDTH. The objective of the paper is to compare the survival times of diabetic patients at UDUTH and BDTH.

2.0 Materials and Methods

2.1 Method of Data Collection

The data used were obtained from clinical studies-Usmanu Danfodiyo University Teaching Hospital, (UDUTH),Sokoto and Barau Dikko Specialist Hospital, (BDTH) Kaduna from 2012-2017, the population-based survey of known diabetes prevalence with a subsequent 6-year mortality follow-up. The hospital set-up for the case study is tertiary health centres in which the medical facilities for the treatment of diabetes are the same. A cohort of 2,613 diabetic patients' folder was observed on 11th April 2018 from two different sources: UDUTH, Sokoto and BDTH, Kaduna. As at the date stated, 1,193 patients in the entire cohort attended UDUTH while the remaining attended BDTH. The life status of the diabetic cohort was ascertained on 11th April 2018. Time from diagnosis of the disease to death defines the failure time while those whose records read "alive" were right-censored because such patients had not died as at the time of the study.

2.2 Statistical Analysis

In this study, Kaplan-Meier Product Limit Estimation Function (Univariate survival analysis) was employed to analyze clinical data of diabetic patients and also we obtained demographic characteristics of diabetic patients. We compare the survival times of diabetic patients attending diabetes clinic at Usmanu Danfodiyo University Teaching Hospital, Sokoto (UDUTH) and Barau Dikko Teaching Hospital, Kaduna (BDTH). All analyses were performed using SPSS (version 20.0).

2.3 Model Specification

2.3.1 Kaplan-Meier Method

If the exact times when deaths occur are known, survival probabilities can be estimated immediately after each individual death without any need to aggregate the data into intervals of one year (or of any other length). This method of estimating the cumulative survival probabilities is called Kaplan-Meier method and it is preferred approach whenever event and censoring times are available [6].

Similarly, to the life-table survival curve, the Kaplan-Meier estimates can be used to plot cumulative survival probabilities. In this instance, however, the plot is in the form of a stepped line, rather than a smooth curve, since the cumulative survival drops at the precise time that a death occurs and remains at a plateau between successive death times.

2.3.2 Estimation Specification

Suppose that the data collected on *n* subjects are denoted by (t, j, c_j), where t is the failure time, c_j is the censoring indicator such that 0 = uncensored case, terminal case and 1 = censored.

Kaplan-Meier product limit estimation is a non- parametric estimation of survival data. The estimation is of the form as proposed by [6]**.**

$$S(t)=\prod_{j=1}^{u}\left[\frac{n-j}{n-j+1}\right]^{c_j} \quad \text{----------(2.1)}$$

Where

S (*t*) = estimated survival function at time t.

$\prod_{j=1}^{u}$ **= Denote the multiplication of the survival times across all cases less than or equal to t.**

t = Time e.g days, weeks, month's year's e.t.c

n = Total number of cases in the sample

j = The number of cases surviving up to time t.

c_j = a constant such that

0 = uncensored case, terminal case

1 = censored

$$SE(t_k)=S(t_k)=\sqrt{\sum\left[\frac{d_j}{n_j(n_j-d_j)}\right]} \quad \text{----------(2.2)}$$

***SE*(t_k) = Estimated cumulative probability at (t) of the event (k)**

d_j = Number of events at time t.

n_j = Number of cases surviving prior to time t. cases not terminated or censored [6]**.**

Table 1: Main clinical characteristics of the diabetic patients.

	Diabetic patients attending UDUTH, Sokoto	Diabetic patients attending BDTH, Kaduna
Cases	1,193	1,420
(Male)	642 (53.8)	765 (53.9)
Sex :		
(Female)	551 (46.2)	655 (46.1)
Age (mean ± SD)	56.9 ± 17.3	52.1 ± 21.7
Age (Group)		
0-19	35 (2.9)	157 (11.1)
20-39	155 (13.0)	195 (13.7)
40-59	461 (38.7)	526 (37.0)
60-79	405 (33.9)	395 (27.8)
≥ 80	137 (11.5)	147 (10.4)
Severity of illness		
Stage1	159 (13.3)	199 (14.0)
Stage2	397 (33.3)	475 (33.5)
Stage3	495 (41.5)	576 (40.6)
Stage4	142 (11.9)	170 (11.9)

Data are mean ± SD for continuous variable and absolute frequency (percent frequency) for categorical variables.

3.0 Results

The main clinical characteristics of the cohort under study are summarized in table 1. With respect to patients that attended UDUTH, patients were slightly older than those attended BDTH (p<0.000). From table 2, by the end of 6-year follow-up, 504 diabetic patients were deceased. Of these, 210 were among the 1,193 patients attending UDUTH (17.6%) and 294 were among the 1,420 attendees' at BDTH (20.7%).

Our interest is only in the variable attendance center, using the Kaplan-Meier product limit estimation function, the mean survival time of attendees at UDUTH and BDTH is 59.913 and 59.812 months respectively whilst their median survival time is 56.000 and 78.000 months respectively (table 3). However, on the overall comparisons, we obtained P-value 0.035 and this reveals that diabetic patients at UDUTH had a better survival than their counterparts at BDTH (table 4).

The difference in survival as a function of attendance at the diabetes clinic was well appreciable from Kaplan-Meier curves (Fig.1 and Fig.2). Patients attending BDTH and UDUTH belonging to stage 1 of illness have 25.1 months and 26.5 months survival time respectively, for stage 2 we have (24.4 months and 24.8 months respectively), for stage 3 (24.9 months and 24.8 months) and

for stage 4 we have 25 months and 24.5 months respectively and this indicates that the mean survival time is slightly higher among the UDUTH attendees' than their counterparts at BDTH especially for stage1 and stage2. Surprisingly, reverse was the case for stage3 and stage4 (Fig. 3).

However, there is significant difference for age and sex among the diabetic patients in their survival times; male patients tend to have lower survival time (58.6 months) than their female counterparts (61.0 months). Similarly, patients of older age have lower survival time than patients of younger age (less than 40 years) and these results were statistically significant (P<0.000).

Table: 2: Case Processing Summary

Case Processing Summary

Attendance centre	Total N	N of Events	Censored	
			N	Percent
UDUTH	1193	210	983	82.4%
BDTH	1420	294	1126	79.3%
Overall	2613	504	2109	80.7%

Table 3: Mean and median survival time of diabetic patients

Means and Medians for Survival Time

Attendance centre	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
UDUTH	59.913	2.265	55.474	64.353	56.000	1.561	52.941	59.059
BDTH	59.812	1.689	56.502	63.123	78.000	.895	53.341	58.657
Overall	60.097	1.352	57.447	62.747	57.000	.634	50.821	55.906

a. Estimation is limited to the largest survival time if it is censored.

Table: 4: Overall comparisons

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	4.463	1	.035

Test of equality of survival distributions for the different levels c
Attendance centre.

4.0 CONCLUSION

From the analysis carried out, we found that survival rate is higher in patients attending a diabetes center located at UDUTH, Sokoto with the mean survival time 59.913 months than their counterparts in BDTH, Kaduna with 59.812 months. The drop line chart shows that the mean survival time is slightly higher among the UDUTH attendees' than their counterparts in BDTH especially for stage1 and stage2. Surprisingly, reverse was the case for stage3 and stage4 (Fig. 3). Patients attending BDTH and UDUTH belonging to stage 1 of illness have 25.1 months and 26.5 months mean survival time respectively, for stage 2

we have (24.4 months and 24.8 months respectively), for stage 3 (24.9 months and 24.8 months respectively and for stage 4, we obtained 25 months and 24.5 months respectively.

Some limitations are to be considered in the interpretation of our results. In particular, because the selection of diabetes clinics under study was not randomized, it is not possible to rule out a self-selection bias. To our knowledge, there are no other reports on the effect of level of health care on survival in diabetic patients.

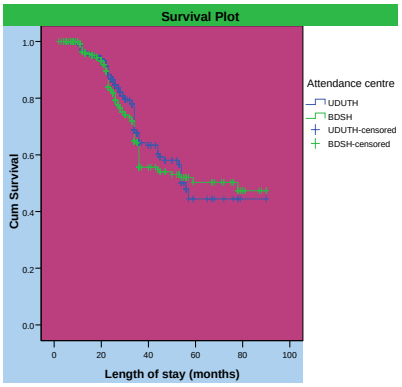

Figure 1: Survival plot for 2,613 diabetic patients data at UDUTH and BDTH from 2012-2017.

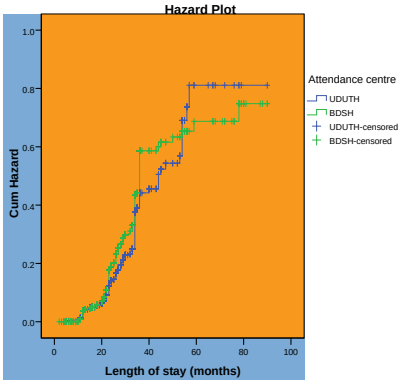

Figure 2: Hazard plot for 2,613 diabetic patients at UDUTH and BDTH from 2012-2017.

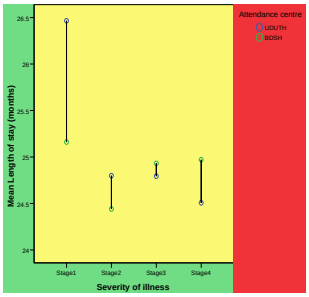

Figure 3. Drop line chart for severity of illness for 2,613 diabetic patients at UDUTH and BDTH from 2012-2017.

ACKNOWLEDGEMENT

Our gratitude goes to the management of Usmanu Danfodiyo University Teaching Hospital (UDUTH), Sokoto and Barau Dikko Teaching Hospital BDTH, Kaduna for the provision of data used for the study.

REFERENCES

[1] Breslow NE, Day NE (1986) Statistical Methods in cancer research: *The Design and Analysis of cohort Studies.*Vol.2 Lyon, France, IARC Scientific, pg. 69-72

[2] Elisa T. L. (1980) *Statistical Methods for survival Data Analysis* lifetime Learning publications.

[3] Moberg E, et al (1993) Estimation of blood-glucose variability in patients with insulin-dependent diabetes mellitus. Scand *J Clin Lab Invest* Vol. 53: 507-514.

[4] www.medgatetoday.com retrieved September 22nd 2018.
[5] Singh B.M, et al (1984) Metabolic control of diabetes in general practice clinics: comparison with a Hospital *clinic.Br Med J* 289:726-728.
[6] Kaplan, E.L. and Meier, P. (1958) Non parametric estimation from incomplete observations *Journal of the American Statistical Association* Vol. 53, 457-48