International Journal of Advanced Statistics and Probability, 3 (1) (2015) 64-71
www.sciencepubco.com/index.php/IJASP
©Science Publishing Corporation
doi: 10.14419/ijasp.v3i1.4032
Research Paper

Analysis of deaths in a Nigerian hospital: the case of Federal Medical Centre (FMC), Umuahia

Ekpenyong, Emmanuel John*, Emenike, Iheanyi Chizaram

Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia state, Nigeria
**Corresponding author E-mail:ekpesstat@yahoo.com*

Copyright © 2015 Ekpenyong, Emmanuel John, Emenike, Iheanyi Chizaram. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

A Poisson regression model was fitted to the causes of death in Federal Medical Center (FMC), Umuahia. The basic violation of the model which is over-dispersion was confirmed absent with the use of a plot of residual against fitted values, the Pearson statistics and the comparison of the log-likelihoods of the Poisson regression and Negative Binomial regression models. Diseases affecting the excretory and neurological systems respectively were found to be insignificant in the model, while the rest of the other causes of deaths – were found to be significant in the model at 5% level of significance. The leading causes of death in the hospital were determined to be diseases affecting the digestive and circulatory system, injuries, endocrine and neonatal related diseases, and cancers.

Keywords:*Poisson Regression; Over-Dispersion; Mortality; Causes of Death; Diseases.*

1. Introduction

Analysis of count data is widely used in medical studies, epidemiology, ecology and many researches of interest. Poisson regression model is one of the most efficient applicable analyses of count data. A good knowledge of what the future holds in stock, no matter how minute has proven pertinent for planning in various areas of life, including the health sector. Fortunately, such knowledge as regards predicting future possible deaths could be obtained with an adequate Poisson Regression model and such other models.

Cause specific mortality is arguably one of the most fundamental metrics of population health. The rates and number of people who die, when, at what age and from what cause is a crucial input into policy debates, planning interventions and prioritizing research for a number of health technologies.

Trends in causes of death provide an important geographical summary of whether society is or is not making progress in reducing the burden of premature (and especially avoidable) mortality and when renewed efforts are needed. If a health information system is not providing timely and accurate information on causes of death by age-sex, major reforms are required to provide health planners with this essential intelligence.

There is the problem of inability to estimate the future number of deaths in Federal Medical Center (FMC), Umuahia, tomake room for effective management and planning. There is also the problem of inability to ascertain the leading causes of death in the hospital as well as the increment or otherwise of recorded deaths.

This research is aimed at making a study of the trend of deaths over the years, to determine the effect of some diseases on the recorded cases of death in Federal Medical Center, FMC Umuahia, and to develop a Poison Regression model that can be used to predict the future number of deaths in Federal Medical Center, FMC, and Umuahia.

Causes of death statistics are the essential data source for monitoring the health of the nations and identifying priorities. They also make an extremely important contribution to the National Burden of Disease Study, which attempts to estimate the causes of loss of healthy life.

Consistent estimates of cause-specific mortality are essential for understanding the overall epidemiological profile of disease in a population. The principal data source for these estimates is civil registration systems. Adequately

functioning systems that produce statistics on causes of death on a regular basis exist in only about one-third of all countries of the world [5].
The optimal source of information in the time of death in populations is vital registration data based on medical certificates of time and cause of death, issued by physicians.
Such information can also be obtained from the medical records of the hospitals. In Nigeria as well as in many developing countries, deaths are usually recorded mainly in the secondary and tertiary hospitals, hence information on deaths are not adequately reported.
Medical records of individuals who die in the hospital may not contain the actual time of death because in several cases, the time documented may be the time of certification of death by the physician. However, the discrepancy is not usually much and more so, the dates of death are usually recorded appropriately, [6]
According to [3], in Sub-Saharan Africa, very little information has been available on cause-specific mortality, let alone data from civil registration systems. Estimates at that time were derived largely from independent disease-specific epidemiological studies and were not examined within the context of an overall demographic "envelope" of mortality, as is required to ensure that claims about causes of death are not exaggerated.
Current evidence on the root-causes of deaths among children younger than 5years is critical to direct international efforts to improve child survival, focused on health promotion and achieved Millennium Development Goal 4. Statistics detailing infant and child mortality, especially in sub-Saharan Africa (SSA) are staggering. Though child mortality has been declining worldwide as a result of socioeconomic development and implementation of child survival interventions, yet 8_8 million children die every year before their fifth birthday.
Globally, substantial headway has been made towards the achievement of MDG4 as the number of under-fives deaths has declined from more than 12 million in 1990 to 7.6 million in 2010. In SSA, the average annual rate of reduction in under-five mortality has accelerated, doubling from 1990-2000 to 2000-2010.
UNICEF also reports that approximately 50% of under-five deaths occur in only five countries, including India, Nigeria, Democratic Republic of Congo, Pakistan and China, that India (22%) and Nigeria (11%) together account for a third of all under-five deaths, that 70% of under-five deaths occur within the first year of life and that the highest rates of child mortality are still in SSA [2].
The estimates ofcause specific mortality were essentially derived from linear models relating mortality levels to broad causes of death [7], with very little effort to incorporate other epidemiological information. It was stated that the large proportion of deaths lumped under other causes reflected the weakness of the estimates, and suggested great caution in interpreting or using them [4].
Smart and Abena [9] used Poisson's regression model to study the incidence of maternal deaths at Okonfo Anokyer Teaching Hospital- KUMASI, Ghana. Analyses were based on data available at the obstetrics and gynecology directorate of the hospital for the period 2000-2010. The data covers all recorded deliveries and maternal deaths for the period January 2000 to December 2010. Since maternal deaths are considered a count data, a Poisson regression model was applied with the years considered as covariates. In all, 132 monthly observations were used in the 11-yearperiod with maternal death cases as the response variable and time (in years) as the predictor variable. The results showed that there was a statistically significant maternal mortality incidence between years 2004, 2005 and 2008 relative to year 2010. A similar approach was used in this study to model the number of deaths in Federal Medical Center (FMC), Umuahia, where the number of recorded death cases was used as the dependent variable and the various causes of death as the independent variables.
Poisson regression model was also used by [8] to model calls to a software support help desk. In his research, the company type and the number of employees which are possible factors that could affect the number of calls were used as the independent variables, whereas the number of calls to the software support help desk was the dependent variable which could be estimated from the model.

2. Methods of data analysis

The data used in this study is a secondary data collected from Federal Medical Center (FMC), Umuahia on the reported cases of death of individuals and their diagnosis.
The data obtained for this research purpose are analyzed using descriptive statistics and Poisson regression model.
In this study, a simple bar graph and tables are used to determine the principal causes of death.
Also a line graph is used to study the trend of recorded deaths in Federal Medical Center (FMC), Umuahia over the years.

2.1. Poisson regression model

Considering the fact that the data set deals with the number of deaths, it becomes paramount to analyze using Poisson's regression model since Poisson's regression is applied where the response variable is a count. Another reason for the application of Poisson's regression is that the number of deaths, which is the response variable is assumed to follow the

Poisson distribution. Typically, the log of the expected value has a linear relationship with the predictor variable(s). The Poisson regression model for the purpose of this research is given as:

$$\log(Y) = \beta_0 + \beta_1 res + \beta_2 \text{cir} + \beta_3 \text{dig} - \beta_4 \text{inf} + \beta_5 \text{inj} + \beta_6 \text{end} + \beta_7 \text{neo} - \beta_8 \text{nut} - \beta_9 \text{hae} - \beta_{10} \text{sur} - \beta_{11} \text{mat} - \beta_{12} \text{gyn} + \beta_{13} \text{can} - \beta_{14} \text{ctd} + \beta_{15} \text{bid} \quad (1)$$

Where Y = the expected number of deaths.
β_0 = the intercept, and β1, β2… β15, are unknown parameters that are estimated from a set of data. These are all obtained using the statistical software SPSS version 16.0. The independent variables are defined as follows:

1) res - RESPIRATORY DISEASES
2) exc - EXCRETORY DISEASES
3) cir - CIRCULATORY/CARDIOVASCULAR DISEASES
4) neu - NEUROLOGICAL DISEASE
5) dig - DIGESTIVE DISEAES
6) inf - INFECTIONS
7) inj - INJURIES
8) end - ENDOCRINE DISEASES
9) neo - NEONATAL DISEASES
10) nut - MALNUTRITION
11) hae - HAEMOTOLOGICAL DISEASES
12) sur - SURGICAL COMPLICATION
13) mat - MATERNAL COMPLICATIONS
14) gyn - GYNOCOLOGICAL DISEASES
15) can - CANCERS
16) ctd - CONNECTIVE TISSUE DISEASE
17) bid - (BROUGHT IN DEAD)
18) non (NOT STATED)

The regression coefficients are estimated using the method of maximum likelihood. The logarithm of the likelihood function is.

$$\ln[L(y,\beta)] = \sum_{i=1}^{n} y_i \ln[t_i \lambda(x_i'\beta) - \sum_{i=1}^{n} t_i \lambda(x_i'\beta) - \sum_{i=1}^{n} \ln(y_i!) \quad (2)$$

The likelihood equations may be formed by taking the derivatives with respect to each regression coefficient and setting the result equal to zero. Doing this leads to set of non-linear equations that admits no closed-formsolution. Thus, an iterative algorithm must be used to find the set of regression coefficients that maximizes the log likelihood.
These values are obtained with the statistical software SPSS version 16.0.

2.2. Assumptions of poisson regression model

The assumptions underlying the use of a Poisson regression model are stated as follows:

i) The dependent variable is a count data.
ii) The dependent variable is theoretically assumed to follow a Poisson distribution.
iii) There is absence of over-dispersion.

Over-dispersion is the most common difficulty encountered in the Poisson regression model. This is typically caused by a Poisson regression model that is missing an important independent variable. In case of over-dispersion, the first remedial approach of the modeler should be to seek more independent variable relevant to predicting the dependent variable.
Outlined below are the various available methods of testing for over-dispersion:
As in any regression analysis, a complete residual analysis should be employed. This involves plotting the residuals against various other quantities such as the regressor variables (to check for outliers and curvature) and the response variable. Various residuals may be of interest, but use is made of the deviance residual which is given below:
The formula for the deviance residual is:

$$d_i = (y_i - \hat{y}_i)\sqrt{2\left\{y_i - \ln\left(\frac{y_i}{\hat{y}_i}\right) - (y_i - \hat{y}_i)\right\}} \quad (3)$$

This will as well be obtained using statistical software SPSS version 16.0.

As an informal model check, a plot of the deviance residuals versus the estimated linear predictors (or fitted values) is made to check for outliers and curvature.

2.3. Pearson statistics and the deviance

Overall performance of the Poisson regression model is measured by two chi- square tests. This is the Pearson statistic

$$P_p = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{\hat{y}_i} \tag{4}$$

And the Deviance

$$D_p = \sum_{i=1}^{n} \left\{ y_i \ln\left(\frac{y_i}{\hat{y}_i}\right) - (y_i - \hat{y}_i) \right\} \tag{5}$$

A value near one (1) for the both statistics confirms the absence of over-dispersion otherwise over-dispersion is present. These values will be obtained using SPSS version 16.0.

2.4. Comparison between poisson regression and negative binomial regression

It is customary to use the negative binominal regression when there is a presence of over-dispersion in the data. Hence one suggested formal test to determine whether there is over-dispersion is to compare the log-likelihood of the Poisson regression and the negative binomial regression. If the log-likelihood of the negative binomial regression is less than that of Poisson's regression, then it offers no improvement, hence over-dispersion is absent and otherwise there is a presence of over-dispersion. This is achieved using SPSS version 16.0.

3. Data analysis

In this section, the methods of analysis discussed above are applied. Firstly, a trend study of death distribution over the years is carried and is shown in figure 1.

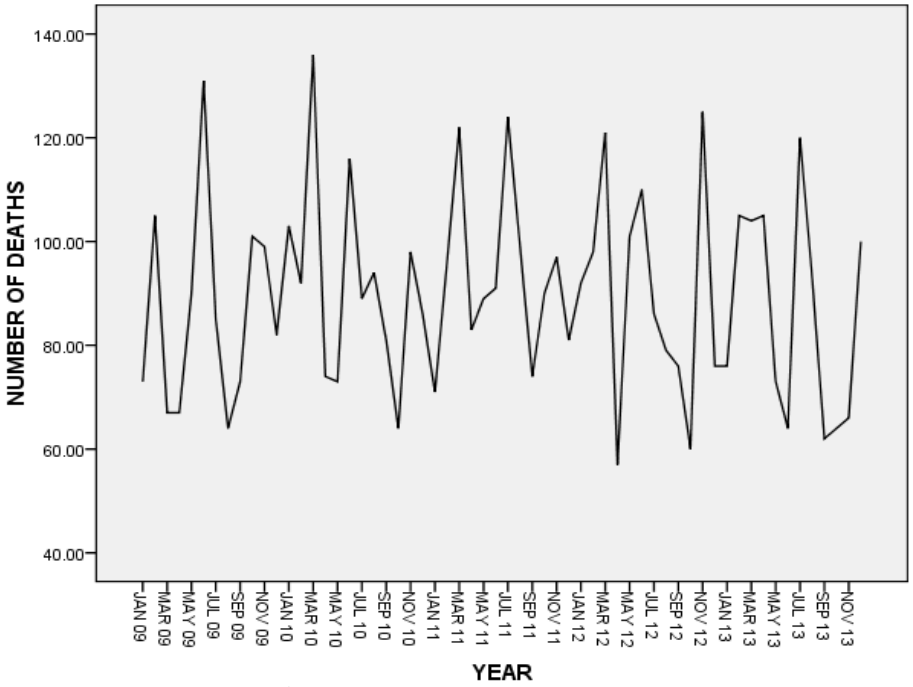


Fig. 1: Trend of Recorded Deaths in Federal Medical Centre, Umuahia.

From Fig. 1 here appears to be no trend over the years.
Over-dispersion which is a major difficulty in the Poisson regression model is tested for its absence. For the sake of accuracy, the following analysis isdone using SPSS version 16.0.

The value for the deviance and Pearson statistics are discovered to be 0.867 for deviance and 0.835 for Pearson chi-square which are close to 1 and is a clear indication that over - dispersion is absent; hence the Poisson regression model is an appropriate model to use.
The value of the log-likelihood is observed for both the Poisson and Negative Binomial Regression models from the goodness of fit table. The log-likelihood for the Poisson regression is observed to be -256.162 while that of the Negative binomial regression is -383.485. This is actually smaller than the log-likelihood for the Poisson regression which indicates that the negative binomial regression does not offer an improvement over the Poisson regression, hence we can infer that over-dispersion is absent from the data.

3.1. Model estimation

Having confirmed the absence of over-dispersion, the regression coefficients are obtained using SPSS Version 16.0 as in table 1.

Table 1: Test for Significance of Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	3.478	.0718	3.337	3.619	2346.128	1	.000
[DISEASE.CLFY=1.00]	-.929	.1349	-1.193	-.664	47.368	1	.000
[DISEASE.CLFY=2.00]	.185	.0972	-.005	.376	3.641	1	.056
[DISEASE.CLFY=3.00]	1.645	.0784	1.491	1.798	439.652	1	.000
[DISEASE.CLFY=4.00]	-.104	.1043	-.308	.100	.994	1	.319
[DISEASE.CLFY=5.00]	2.105	.0761	1.956	2.254	765.656	1	.000
[DISEASE.CLFY=6.00]	-.236	.1081	-.447	-.024	4.750	1	.029
[DISEASE.CLFY=7.00]	1.013	.0838	.848	1.177	145.903	1	.000
[DISEASE.CLFY=8.00]	.496	.0911	.317	.674	29.640	1	.000
[DISEASE.CLFY=9.00]	.837	.0860	.668	1.005	94.723	1	.000
[DISEASE.CLFY=10.00]	-4.394	.6502	-5.669	-3.120	45.672	1	.000
[DISEASE.CLFY=11.00]	-1.027	.1399	-1.301	-.753	53.942	1	.000
[DISEASE.CLFY=12.00]	-.757	.1271	-1.006	-.508	35.475	1	.000
[DISEASE.CLFY=13.00]	-1.080	.1426	-1.360	-.801	57.361	1	.000
[DISEASE.CLFY=14.00]	-3.008	.3310	-3.657	-2.359	82.584	1	.000
[DISEASE.CLFY=15.00]	.274	.0953	.087	.460	8.252	1	.004
[DISEASE.CLFY=16.00]	-2.143	.2216	-2.578	-1.709	93.505	1	.000
[DISEASE.CLFY=17.00]	1.816	.0774	1.664	1.968	550.184	1	.000
[DISEASE.CLFY=18.00]	0a	.	.	.	.	.	.
(Scale)	.835b						

\Dependent Variable: NO.OF.DEATHS
Model: (Intercept), DISEASE.CLFY

a. Set to zero because this parameter is redundant.

b. Computed based on the Pearson chi-square.

The second column of Table 1, (column 'B') is the corresponding regression coefficients. It is also observed that variables 2 (expected number of deaths due to excretory diseases) and 4 (expected number of deaths due to neurological diseases) are statistically insignificant in the model at 5% level of significance, as their significant probability values were greater than 0.05 ($p > 0.05$) and were therefore, inappropriate to be included in the model. All other variables are significant in the model at 5% level of significance.

3.2. Plot of residual

This method is the informal check, which involves the plot of the deviance residuals versus the predicted values of the mean of response. These values and the resultant plot were obtained using SPSS version 16.0. The plot is presented in Figure 2.

Deviance Residual against fitted values.

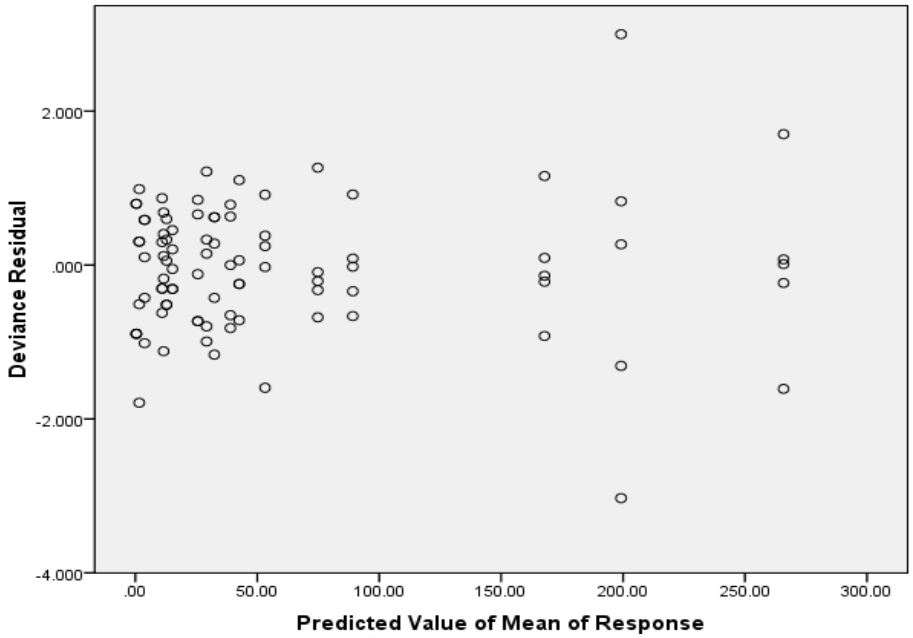


Fig. 2: Plot of Residual versus Predicted Values.

The resulting scatter plot does not have any outlying point which is one possible cause for concern and also, there is no curvature which suggests that there is no over-dispersion in the data set.

3.3. Leading specific causes of death

A summary of the leading specific causes of death is made below with the aid of a simple bar graph of the causes of death against frequencies as shown below:

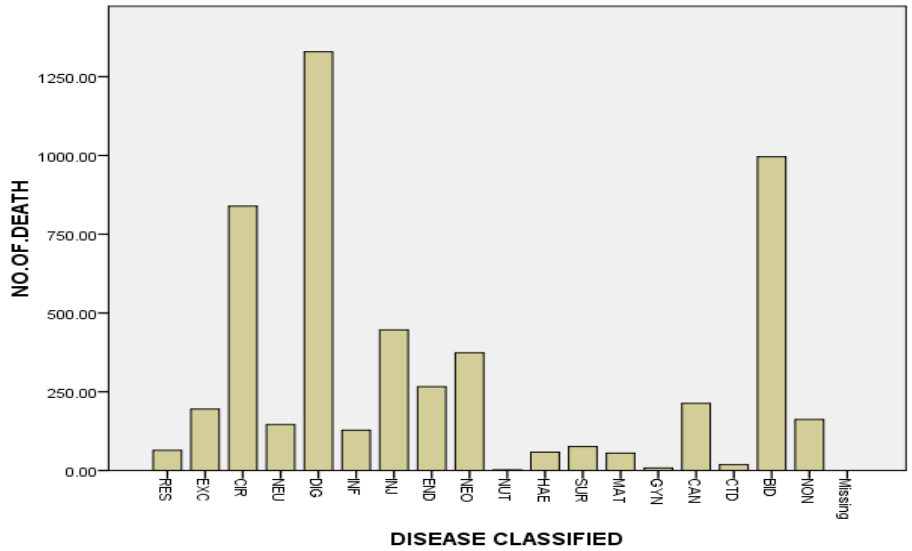


Fig. 3: A Simple Bar Chart of Specific Causes of Death.

Fig. 3 and Table 2 indicate the descriptive representation of the level of contributions of the classified diseases to death.It should be noted that the causes of death indicated here are not individual causes of death, but a classification of some causes.

The major causes of death with their percentage causes as derived from Figure 3 and Table 2 are given in Table 3 and they account for just over 80 percent of all the deaths.

Table 2: Percentage contributions of Causes of Death

Disease Classified	Number of Deaths	Percent (%)
res	64	1.2
exc	195	3.6
cir	839	15.6
neu	146	2.7
dig	1329	24.7
inf	128	2.4
inj	446	8.5
end	266	4.9
neo	374	7.0
nut	2	0.0
hae	58	1.1
sur	76	1.4
mat	55	1.0
gyn	8	0.1
can	213	4.0
ctd	19	0.4
bid	996	18.5
non	162	3.0

Table 3: Major causes of death and their percentage contributions

Disease Classified	Number of Death	Percent
dig	1329	24.70%
bid	996	18.50%
cir	839	15.60%
inj	446	8.30%
neo	374	7.30%
end	266	4.90%
can	213	4%
Major Causes	4463	83.30%
All Causes	5376	100.00%

Diseases affecting the digestive system accounted for 24.7% of total deaths. 18.5% of the total deaths were people brought in dead (BID) and 15.6% of the total deaths is caused by circulatory system diseases, while injury-related, neonatal, endocrinal and cancer-relateddiseases cause 8.3%, 7.3%, 4.9%, and 4% respectively. So the result suggests that diseases affecting the digestive and circulatory system, injuries, neonatal and endocrinal diseases, and cancers are the leading causes of death.

4. Discussion of results

Disease affecting the respiratory system was found to among the leading causes of death in UK [5]; unlike that of FMC Umuahia where diseases affecting the digestive system were the leading cause of death. This could have been informed due to massive increases in tobacco use in UK as compared to Umuahia. Also, this is true of Umuahia, Abia State, vis-à-vis Nigeria because of the type of foods we eat and our feeding habits.

According to Aryal et al [1], the three important causes of mortality were COPD, CVA and Cirrhosis of liver in the NMC teaching hospital. Annual health report 2002/2003 also stated that COPD and CVA were top two cases of mortality in Bir hospital2. Similarly, Tropical Hospital at Kathmandu also reported 11. 11percent inpatient's deaths were due to COPD2. The ten percent hospital death was due to cancer under the study. Bhaktapur Cancer Hospital reported out of 4389 inpatients 170 patients died due to different types of cancer from the year 1999 to 2003. The death rate was decreased from 5.09 percent in the year 1999 to 2.83 percent in the year 2003. Most of these causes are amenable to effective interventions, hence are subjects of research and action orientated programs.

5. Conclusion

It is observed from the study that there is no trend in the number of recorded deaths in Federal Medical Center (FMC), Umuahia over the years 2009-2013 which suggest that the number of recorded deaths has not changed significantly. However, the graph showed some spikes, which suggests fluctuation in the number of recorded deaths in the hospital for each observed months

Furthermore, a Poisson regression model was fit, which could be used to predict the future number of deaths in Federal Medical Center (FMC), Umuahia. In the model, the number of recorded death cases was used as the dependent variable

while the causes of death were used as the independent variables. Such information would inform the entire public on the impending danger as it regards loss of life.

The major causes of death were identified to be diseases affecting the digestive and circulatory system, injuries, endocrine and neonatal causes and cancers which will inform the prioritization of preventive strategies to caution the effect of the principal causes.

A large proportion of the recorded deaths (18.5%) in the hospital were people brought in dead, which could be informed by peoples' nonchalant attitude towards receiving professional medical attention.

References

[1] U. L Aryal, P. Dhakal, S. Shrestha, S. Upadhyaya and M. R. Chhetri: A Statistical Analysis of Patient Death in NMC Teaching Hospital: *Journal of Nepal Health Research Council, 4(2), (2014), 41-44.*

[2] M. A Bambgoye, O. C Cecilia, E. Adejuwonlo and D. Duro: A hospital based estimate of major causes of death among under-five children from a health facility in Lagos, southwest Nigeria: possible indicators of health inequality. *International Journal for Equity in Health, 11(39) (*2012), *1-15*.

[3] R. G Feachem and D. T Jamison: Disease and Mortality in Sub-Saharan Africa: Oxford University Press, New York, 1991.

[4] R. G. Feachem, D. T. Jamison and E. R. Bos: Changing Pattern Disease and Mortality in Sub-Saharan Africa: New York University Press. New York, 1993.

[5] A. D. Lopez, O. B. Ahmad, M. Guillot, B. D. Ferguson, J.A. Salomon, C. J. L. Murray and K. H. Hill : World Mortality in 2000: Life Tables for 191 Countries: WHO. Geneva. 2002

[6] B. O. Nwosu, N.O Eke, O.J. Osakwe, C.O. Eke and N.P. Obi. Weekend versus Weekday hospital deaths: Analysis of in-patient data in a Nigerian tertiary healthcare center. *Nigerian Journal of Clinical Practice, 16(4),* (2014), *501-504.http://dx.doi.org/10.4103/1119-3077.116900*.

[7] S. H. Preston. Mortality Patterns in National Populations: With Special Reference to Recorded Causes of Death. Academic Press. New York, 1976

[8] L. Russ: An animated guide: An Introduction to Poisson Regression. K&L consulting services, King of Prussia, PA, USA, 2010.

[9] A. S. Smart and K. B. Abena: Poisson Regression Modeling for Incidence of Maternal Deaths in Ghana. *Mathematical Theory and Modeling, 3(2),(2013),30-37.*