STATISTICAL ANALYSIS OF NON-COMMUNICABLE DISEASES IN A HEALTH FACILITY IN TAKORADI, GHANA

Michael Kwabena Asirifi[1], Frank B. K. Twenefour[1] Mark Kofi Aremu[2]

[1] Department of Mathematics and Statistics, Takoradi Polytechnic, P. O. Box 256, Takoradi, Ghana
[2]Planning Office, Takoradi Polytechnic, P. O. Box 256, Takoradi, Ghana

ABSTRACT: *Current trends indicate that global Non-Communicable Diseases (NCD) accounts for about 60% of deaths and will increase by 17% over the next 10 years with poor and disadvantaged populations disproportionately affected, widening health disparities between and within countries. It is against these challenges that "Statistical Analysis of Non-Communicable diseases" was undertaken. The main objective of this paper was to determine the age groups that are affected most and also to determine the trend of each of the selected Non-Communicable Diseases. To achieve this, a five-year data set was collected from Takoradi Hospital. Results of the analysis of the data depict that, females dominate those who are suffering from Non-Communicable Diseases. Also, people within 20 to 34 year group are mostly affected by Non-Communicable Diseases. It also reveals that the number of cases of the Non-Communicable Diseases analyzed have declining trend with the exception of anemia. It was therefore recommended that, people from all walks of life must give due consideration to their diet and thus eat balanced diet and do regular exercise to keep them healthy.*

KEYWORDS: Non-Communicable Disease, Time Series, Forecast, Health Facility, Takoradi-Ghana.

INTRODUCTION

The term disease broadly refers to any condition that impairs normal function, and it is therefore associated with dysfunction of normal homeostasis. Commonly, the term disease is used to refer specifically to infectious diseases, which are clinically evident diseases that result from the presence of pathogenic microbial agents, including viruses, bacteria, fungi, protozoa, multicellular organisms, and aberrant proteins known as prions. An infection that does not and will not produce clinically evident impairment of normal functioning, such as the presence of the normal bacteria and yeasts in the gut, or of a passenger virus, is not considered a disease. By contrast, an infection that is asymptomatic during its incubation period, but expected to produce symptoms later, is usually considered a disease. Non-infectious diseases include such diseases as most forms of cancer, heart and genetic diseases (*Bosu*, *2005*).

According to the World Health Organization (2002), "the risk for Non- Communicable Diseases (NCDs) is gaining importance in Africa with a prevalence of high blood pressure estimated at 30% - 40% although prevalence data from national surveys are generally inadequate". The World Health Report 2001 indicates that NCDs, in 1998, accounted for almost 60% of all deaths and 43% of global burden in that year. Seventy-five percent of total deaths resulting from NCDs occur in the developing countries. The treatment of NCDs is known to be expensive, labour-intensive and needs high technology and sophisticated machines/tools which most developing countries are lacking.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

A number of studies also suggest that non-communicable diseases will soon be the most important cause of morbidity and mortality in all developing countries (Agyei-Mensah, 2004; UN, 2005; and Addo, Smeeth & Leon, 2007). Chronic non-communicable diseases are currently the main cause of both disability and death worldwide (WHO 2002). The expectation is that the proportion of chronic diseases would be comparatively greater in developing countries in some years to come (Bela and Prashant, 2005). This is because developing countries have a greater stock of mortality and morbidity from infectious diseases that are curable by improvements in nutrition and sanitation standards through relatively simple medical means. It is also expected that there would be disparities between male and female chronic diseases burden due to biological make up and gender roles associated with African communities. There has been a projection that about 28 million people in sub-Saharan Africa will die from a chronic disease over the next 10 years (WHO, 2002). According to the projections, the number of elderly people in the world will exceed the number of young people by the middle of this century.

In African countries, ageing-related issues have low priority in many governmental sectors. This low priority may be due to inadequate information relevant to the population and its situation (WHO, 2002). Ageing of the population poses significant problems to the healthcare system, including increased costs, inappropriate modes of service delivery and problems of chronic disease prevention. Agyei-Mensah (2004) stated that healthy ageing is largely determined on social, economic opportunities, health status and access to health care over the life course. Worldwide, older persons carry a greater share of the burden of chronic diseases. According to Bela and Prashant (2005), some 63% of older people still find it hard to access health care and 72% do not have enough income to access basic services such as water, electricity, food, and decent housing. In such situations, Non-Communicable Diseases tend to be complex conditions especially with advances in age; they are often long-lasting and persistent in their effects and can produce a range of complications. The gains in life expectancy as well as improvement in medical advancement, sanitation and nutrition in most developing countries has resulted in the progression of chronic illness especially hypertension (Bosu 2007).

METHODS

Time series

A time series is a collection of observations on a variable (quantitative) of interest over a period of time, usually taken at equal intervals of time. Quarterly sales at a retail store over the last four years and monthly electricity consumption of a home over the last two years are example of time series data (DeLurgio, 1998, Keller and Warrack, 2003).

Components of Time Series

A time series is made up of some or all of the following four components (DeLurgio, 1998; Keller and Warrack, 2003):

i. Secular Trend
ii. Cyclical Variation
iii. Seasonal Variation
iv. Irregular fluctuations

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

Secular Trend is the overall long-term pattern in the values of the series. It is about whether the values in the series are increasing or decreasing over time or the values are neither increasing nor increasing or the values are declining after a time of increase. Plotting the series against time will usually reveal the trend. Employing one of several smoothing techniques may also help to identify the trend.

Cyclical Variation refers to the up and down movement in the values of the series about the trend line that repeat themselves after a period of more than one year; the whether situation of experiencing severe drought every 10 or so years (Elniño) is an example of cyclical variation.

Seasonal Variations are fluctuations in the values of the data that usually occur within a short period of time and repeat themselves at the same point in the adjoining time period. The period is usually a year or less as opposed to the long-term nature of cyclical variations. The high volume of sales around Christmas and Easter and the fall in sales immediately after Christmas and Easter are examples of seasonal phenomena.

Irregular Fluctuations are sharp changes in the values of the series that are unexplained or unpredictable; they present themselves as sharp spikes or falls in the values of the variable in a very short time. For instance, supply of tents may fall sharply due to a strike at the plant making the tent or sales of tents may rise sharply due to high demand for tents after an earthquake.

We note that the cyclical variation is often associated with economic or business data. For this reason, it is sometimes ignored in the analysis of time series that are not clearly economic (DeLurgio, 1998; Keller and Warrack, 2003).

Decomposition of Time Series

Decomposing a time series allows us to know which components are present and gives us the opportunity to have a better understanding of the component of the series, which puts us in a good stead to make future predictions more accurately (Keller and Warrack, 2003).

Choice of Model

A time series Y is said to follow the Additive Model if it is thought to be the sum of some or all of the component parts: Secular Trend (*T*), Cyclical Variation (*C*), Seasonal Variation (*S*), Irregular fluctuations (*I*).

That is $Y = T + C + S + I$, assuming all the components are present.

The series is said to follow the Multiplicative Model if it is thought to be the product of some or all of the component parts. That is $Y = T \cdot C \cdot S \cdot I$ (Bowerman and O'Connell, 1997; DeLurgio, 1998; Harper, 1991).

It is recommended that the additive model should be used when the plot of the series reveals that the trend is relatively flat (Harper, 1991) or the seasonal effect at any point in time is a certain fixed value above or below the combined trend-cycle, but where the seasonal effect is a percentage of the combined trend-cycle then the multiplicative model is appropriate (DeLurgio, 1998). Given that the seasonal effect being a percentage of the combined trend-cycle could result in the seasonal effect at any point in time being a certain fixed value above or below the combined trend-cycle "the multiplicative method is equally able to handle both flat and steep trends, it is clearly the better general purpose method, and if a choice is available

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

the multiplicative method should normally be selected" (Harper, 1991, p.195). Additionally, the additive model assumes that the components are independent which may not be realistic.

Estimation of Components

Fitting the Trend by the Method of Least Squares

The trend in a time series could be linear or curvilinear. A linear (straight line) trend is described by equation $y_t = \beta_0 + \beta_1 t$ and a curvilinear trend by $y_t = \beta_0 + \beta_1 t + \beta_2 t^2$ or $y_t = \beta_0 {\beta_1}^t$ among others, where y_t is the value of the series at a point t in time and the $\beta_i s, i = 1, 2, 3.$, are constants parameters – their values are never known. For any meaningful analysis and forecasting to be done we need to obtain unbiased estimates of the $\beta_i s, i = 1, 2, 3 \ldots n$. The convention is to use the Method of Least Squares to estimate the $\beta_i s$. The procedure has been illustrated by Freund (2001) in Mathematical Statistics (pages 501-505).

Estimating Seasonal and Trend Component

Moving Averages (MA)

Given the series $y_1,\ y_2, \ldots, y_i, \ldots\ y_n$ a moving average of order *k* is defined by the sequence of averages (Spiegel, 1992):

$$(y_1 + y_1 + \ldots + y_k)/k\,, (y_2 + y_3 + \ldots + y_{k+1})/k\,, (y_3 + y_4 + \ldots + y_{k+2})/k, \ldots$$

Moving averages of suitable orders are one of several smoothing techniques employed in time series analysis. They are particularly employed to isolate the seasonal effect in a time series. They remove the short-term seasonal variation and irregular fluctuation so that what is left of the series is a combination of the trend and cycle (Bowerman and O'Connell, 1997; DeLurgio, 1998).

From the perspective of the multiplicative model:

$$MA = TC$$

It follows that

$$Y/(M \cdot A) = Y/(T \cdot C) = S \cdot I$$

The seasonal effect in the form of an index is then found by averaging the *S I* values corresponding to the same time point. For instance if we are dealing with monthly data, then the average of all *S I* values for January gives us the seasonal index for January, the average of all *S I* values for February gives us the seasonal index for February, and so on. If the sum of all the averages (which is usually expressed as a percentage) is not exactly 1200 then they must be adjusted to give a sum of 1200 by multiplying the averages by the adjustment factor, 1200/(Total of Averages) (Bowerman and O'Connell, 1997; DeLurgio, 1998).

Measures of Model Accuracy

$e_t = y_t - \hat{y}_t$ represents the deviation (or error) of the fitted (or forecasted) value $\hat{y}_t$ from the actual value y_t of the series. Several measures of model accuracy based on the difference e_t are presented below. Namely Mean Absolute Deviation (MAD), Mean Square Deviation (MSD),

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

Mean Absolute Percentage Error (MAPE), Standard Error of the Estimate S_{sy} and Adjusted Coefficient of Determination $\overline{R}^2$

MAD, MSD and MAPE are given as:

$$MAD = \frac{\sum_{t=1}^{n} \left| (y_t - \hat{y}_t) \right|}{n}$$

$$MSD = \frac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{n}$$

$$MAPE = \frac{\sum_{t=1}^{n} \left| (y_t - \hat{y}_t)/y_t \right|}{n}$$

Where n is the number of data points.

The closer the values of the actual series to the model values the smaller will be the deviation or error values ($e_t s$) and the better will be the model as a fit or forecasting tool. Thus smaller values of MAD, MSD and MAPE are an indication that the model is a good fit or is good for forecasting. However these statistics are not that meaningful on their own; they are meaningful when comparing one model with another. That is, given a number of models one would choose the one with relatively smaller values of MAD, MSD and $MAPE$ (Bowerman and O'Connell, 1997;

DeLurgio, 1998).

Standard Error of the Estimate S_{xy}

The standard error of the estimate S_{xy} *is given by*

$$S_{xy} = \sqrt{\frac{SSE}{n-k}}$$

Where n is the number of observations, k is the number of estimated parameters and SSE is the sum of the squares of the error terms or the residuals. S_{xy}, like MAD, MSD and $MAPE$, measures the scatter of the actual values y_t *about the model* $\hat{y}_t$. Thus a model with comparatively smaller value of S_{xy} is a better fit. $S_{xy}{}^2$ measures the variability in the original data that is not accounted for by adopting the model $\hat{y}_t$ (DeLurgio, 1998; Keller and Warrack, 2003).

Adusted Coefficient of Determination $\overline{R}^2$

The adjusted coefficient of determination $\overline{R}^2$ is given by

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$\overline{R}^2 = 1 - \frac{S_{xy}{}^2}{s_y^2}$$

The quotient S_{sy}^2 / S_y^2 measures the proportion of the variability in the actual data that is unexplained by adopting the model $\hat{y}_t$. It follows that $\overline{R}^2$ measures the proportion of the variability in the original data is accounted for by the model $\hat{y}_t$ (DeLurgio, 1998; Keller and Warrack, 2003).

RESULTS

An examination of data in table 1 and 2 shows that Anaemia, Hypertension, Diabetes, Asthma and Eye related problems featured in the list of the top 5 non-communicable diseases from the year 2008 through to 2012. It can be observed from table 1 that there were more reported cases of non-communicable diseases by females compared to that of their male counterparts in all the years except for Asthma where reported cases of males and females were at par.

Table 1: Number of Patients for both sex and its percentages for the selected diseases from 2008 to 2012

Year	Anaemia		Hypertension		Diabetes		Asthma		Eye	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
2008	162	271	1068	1923	150	373	106	113	5301	6626
2009	257	325	802	1515	161	246	163	206	8497	9541
2010	203	255	352	611	84	145	164	180	5166	7199
2011	408	569	232	375	77	114	111	63	5670	7459
2012	938	1470	220	279	51	92	97	78	4152	5944
Total	1968	2890	2674	4703	523	970	641	640	29086	36769
Percent	41%	59%	36%	64%	35%	65%	50%	50%	44%	56%

Table 2 presents the total number of patients within each age-group from 2008-2012. It could be seen that for the cases of Anaemia, the age group 30-39 years had a higher percentage (17.68%) of reported cases than that of the various age groups. The age group 50-59 recorded the highest number of cases that is 27.95% and 25.85% for Hypertension, Diabetes respectively. Also the age group 10-19, recorded 22.87% and 21.68% for Asthma and Eye respectively.

Table 2: Total number of patients within each age-group from 2008 to 2012

Age	Anaemia		Hypertension		Diabetes		Asthma		Eye	
Group	Total	%	Total	%	Total	%	Total	%	Total	%
<1	195	4.01	0	0.00	0	0.00	10	0.78	2211	3.36
1-9	782	16.10	3	0.04	5	0.33	151	11.79	9141	13.88
10-19	816	16.80	52	0.70	28	1.87	293	22.87	14280	21.68
20-29	766	15.76	567	7.70	40	2.68	217	16.94	5165	7.84
30-39	859	17.68	750	10.17	130	8.71	190	14.83	7231	10.98
40-49	345	7.10	1002	13.60	325	21.76	136	10.62	8265	12.55
50-59	457	9.41	2061	27.95	386	25.85	149	11.63	6602	10.03
60-69	302	6.22	1488	20.18	322	21.56	76	5.93	6030	9.16
$\geq$ 70	336	6.92	1450	19.66	257	17.21	59	4.61	6930	10.52

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

Total	**4858**	**100**	**7373**	**100**	**1493**	**100**	**1281**	**100**	**65855**	**100**

The legend of figure1 describes three items (actual, fits and forecast), the actual values represents the values of anaemia patients recorded. The fitted values represent the number of anaemia patients suggested by the fitted model and the forecast values are the number of anaemia cases projected to be recorded after 2013.

The trends in the actual and the fitted values suggest there is an increasing trend in the number of anaemia patients reported at the various health facilities. The Forecasts line also shows the number of anaemia patients to be expected from 2013 to 2017. The number of anaemia patients is projected to increase from 2275 in 2013 to 4013 in 2017 (see Figure 1).

From Figure 2 it can be seen that the actual line descends from the left corner. This shows a decreasing trend in the number of hypertension patients. Forecast line shows the predicted number of patients to be expected in the future years.

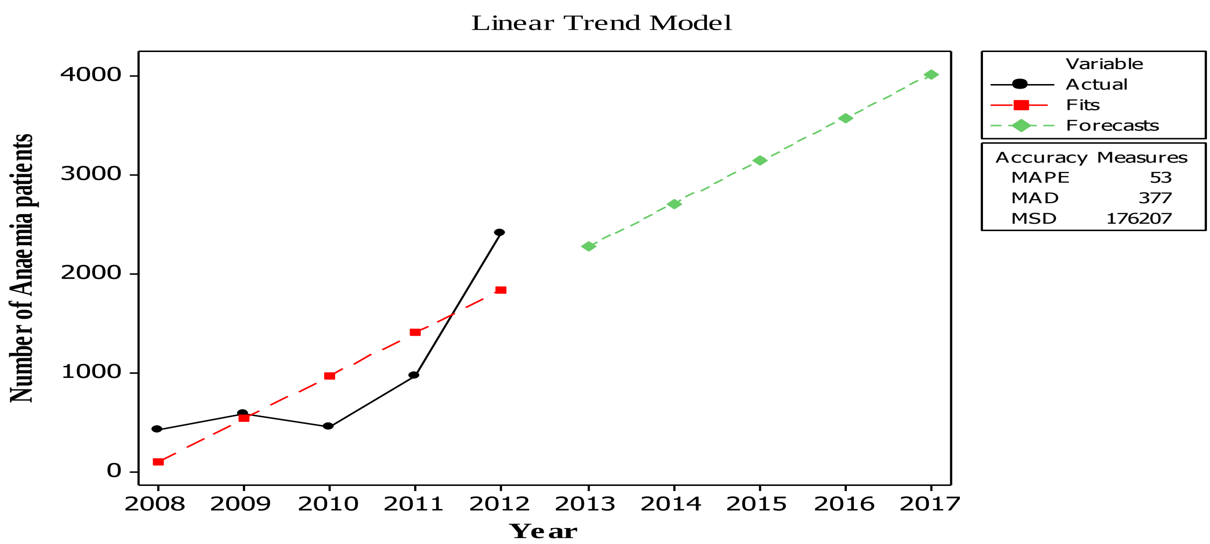

Figure 1: Trend analysis plot for anaemia patients

The trend line depicts that there will be decreasing number of people who will be reporting with cases of hypertension from 2013 onwards.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

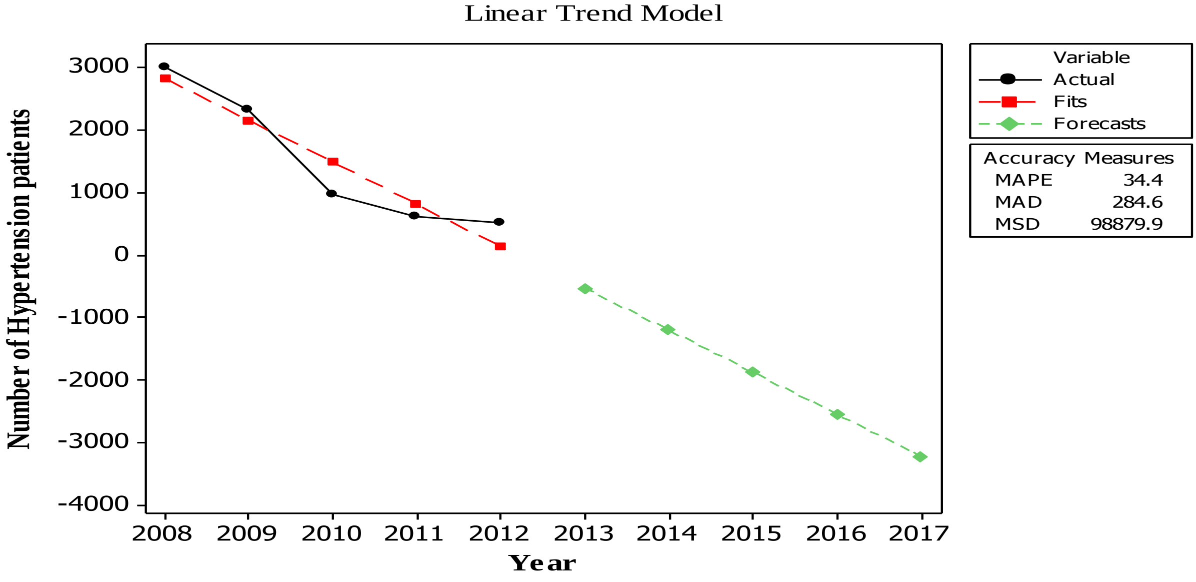

Figure 2: Trend analysis plot for hypertension patients

From figure 3, it can be seen that the actual line descend from the left corner. This shows a decreasing trend in the number of diabetes patients. The forecast line also shows the predicted number of patients to be expected in future years. The number of patients to be expected in 2013 is six (6) and will continue to decrease in the predicted years.

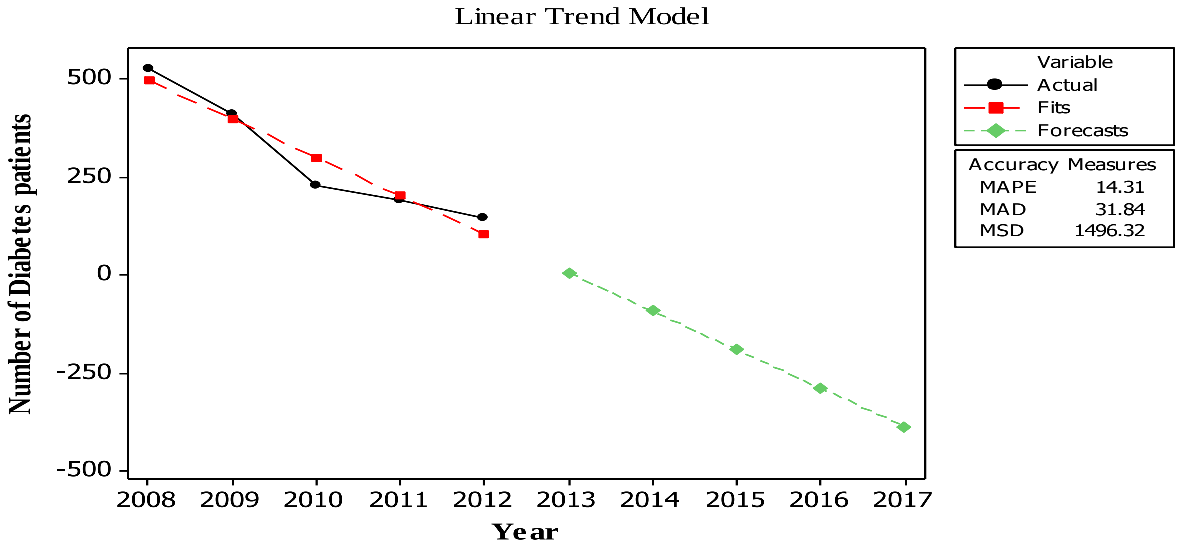

Figure 3: Trend analysis plot for diabetes patients

Figure 4, shows a decreasing trend in the number of asthma patients. This is confirmed by the fits line. Forecast line shows the predicted number of patients to be expected in the future years

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

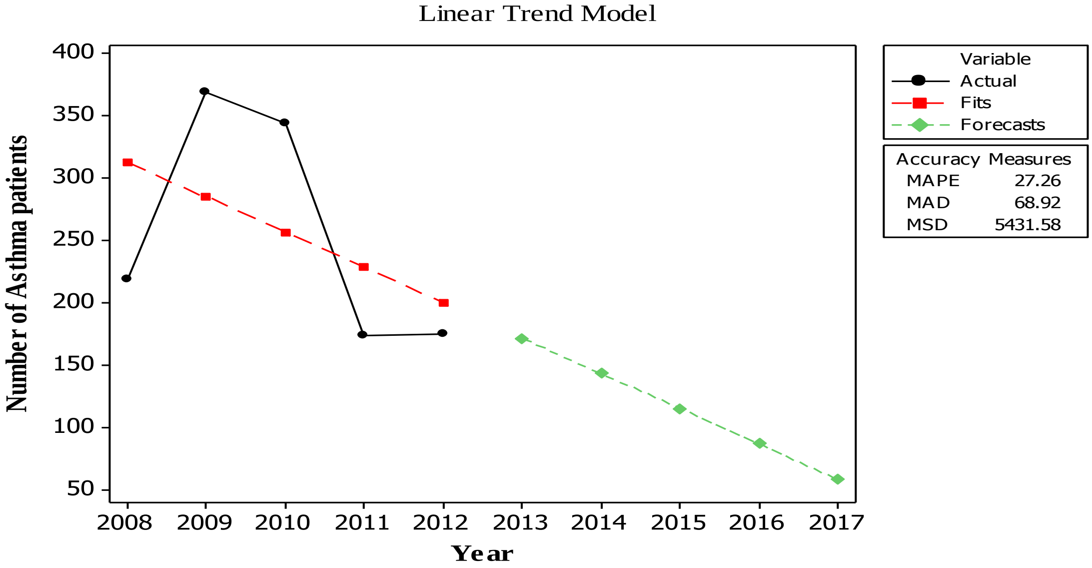

Figure 4: Trend analysis plot for asthma patients

The number of patients expected in 2013, 2014 and 2015 are 171, 143, and 115 respectively. Also eighty-six (86) and fifty-eight (58) patients are expected in 2016 and 2017 respectively.

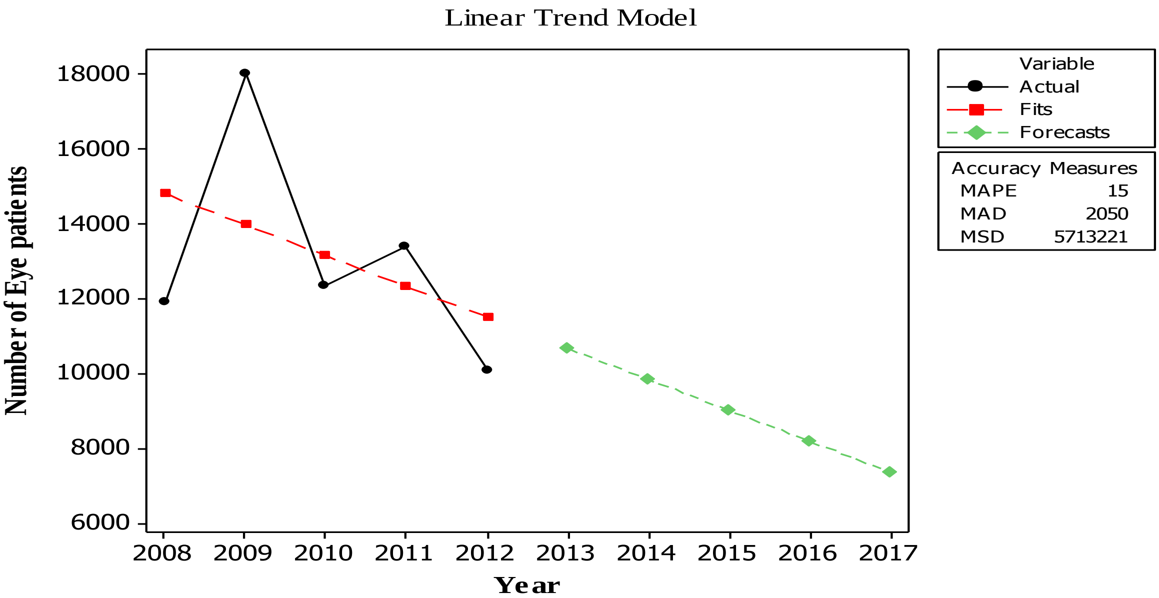

Figure 5: Trend analysis plot for eye patients

Figure 5 presents the Forecast line and shows the predicted number of patients to be expected in the future years. The number of patients expected in 2013, 2014 and 2015 are 10690, 9863 and 9036 respectively. Also 8208 and 7381 patients are expected in 2016, and 2017 respectively.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

DISCUSSION AND CONCLUSION

From our analysis we found out that Anaemia patients recorded at the hospital from year 2008 through to 2012 was 4858. It was observed that female patients were dominant. It was also observed that the leading age group recorded among anaemia patients were those within 20-34 years. Apart from the fact that 20-34 years is the dominant age group, there are increasing rates in the other age groups below 20-34. Also, those less than one year and those within 1-4 years were more than those within the age bracket of 60 years and over. Out of the 7377 hypertension patients recorded from 2008 to 2012, majority (64%) were females while the rest (36%) are males. Thus, females are dominant in terms of hypertension. Over twenty-seven percent (27.94%) hypertension patients were within 50 - 59 years. Also, few Hypertension patients were within 18-19 years. In the case of Diabetes patients, there were 1493 patients from 2008 to 2012, of which 65% were females and 35% were males. Also, it was found that most patients were within 50 - 59 year group constituting 25.85%

The recorded number of Asthma patients from 2008 to 2012 was 1281. This figure was evenly distributed between the males and females comprising 50% each for both sexes of the population of asthmatic patients. Others are the following: 21.86% are within 20-34 year group, 25.50% are within 35-40 year group, and 10.50% are also within 10-14 year. In the population of 65855 eye patients recorded from 2008 to 2012, fifty –six percent (56%) were females while forty-four (44%) were males. The dominant age group was 10-19 years. It was further observed that, there were also patients who were below this age group (10-19 years) suffering from eye related problems as well. It was also observed that, there is an increasing trend in the number of anaemia cases at a slower pace and a decreasing trend in the number of cases of hypertension but at a faster rate than anaemia. The cases recorded for diabetes indicate that this disease is increasing at a very fast pace. However, the rate of increase in both asthma and eye related problems could be said to be at slower rates. The objective of this research was to find the age group affected most by the five diseases analyzed. The study found out that, any patient diagnosed of the five Non-Communicable Diseases considered, is most likely to be within 20-34 years. It was also found that, the female population were the ones most affected by Non-Communicable Diseases as compared to their male counterparts.

The trend analysis also reveals that, there is a decline in the trend of all the Non-Communicable Diseases analyzed except Anaemia. It is also expected that the number of Anaemia patients in 2013 will decline from 2013 to 2017. The number of Diabetic patients expected in 2013 is 6 and will decline throughout the years to 2017. The number of Asthma patients expected in 2013 is 171 and will decrease to fifty-eight (58) patients in 2017. Finally, the number of patients with Eye related problems expected in 2013 is 10690 and this number will decline to 7381 patients in 2017.

Generally, the results of the analysis support the findings of World Health Organization in the literature reviewed. That is, female sex is mostly affected by Non-communicable diseases and mostly found within the range of 13-40 years. This is based on the findings in the analysis that most patients are within 20-34 years and also were found to be mostly females.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

REFERENCES

[1] Addo, J., Smeeth, L. & Leon, D. A. (2007).Non-Communicable Diseases in Sub-Saharan Africa: *A Systematic Review***.** 50 (2), 1012-1018.

[2] Agyei-Mensah, S. (2004).*Accra (Ghana) in Transition: Epidemiological Change at the turn of the 21st century*. Paper presented at the Seminar on Health in African Cities, Harvard Centre for Population Studies, Harvard University.

[3] Bela, S. and Prashant, M. (2005). Risk Factor Surveillance for Non-Communicable Diseases (NCDs): *The Multi-Site ICMR-WHO Collaborative Initiative*. Nos. 25, pp 353.

[4] Bosu, W. K. (2007). Ghana's National NCD Programme*: History*, *Prospects and Challenges.* Paper Resented At The 1st Annual Workshop, British Academy UK-Africa Academic Partnership On Chronic Disease In Africa, Noguchi Memorial Institute For Medical Research.

[5] Bowerman, B. L. and O'Connel, R. T. (1997), *Applied Statistics: Improving Business Processes.* Bonston, USA: Irwin/Mcgraw-Hill.

[6] DeLurgio, S. A. (1998), *Forecasting Principles and Applications*. 1st ed. Bonston, USA: Irwin/Mcgraw-Hill.

[7] Freund, J. E. (2001), *Mathematical Statistics*. 5thed. New Delhi: Prentice-Hall.

[8] Gordor, B. K. and Howard, N. K. (2000). *Elements of Statistical Analysis*. The City Printers, Accra, Ghana.

[9] Harper, W. M. (1991), *Statistics*. 6th ed. London, UK: Longman Group Ltd.

[10] Keller, G and Warrack, B. (2003), *Statistics for Management and Economics*. 6thed. Madrid, Spain: Thompson, Brooks/Cole.

[11] Spiegel, M. R. (1992), *Statistics*. 2nd ed. London, UK: McGraw-Hill International (UK) Ltd

[12] World Health Organization. (2002). Joint WHO/FAO Expert Consultation on Diet, Nutrition and the Prevention of Chronic Diseases. Report No. 916.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)