ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.4, No.13, 2014

Analysis of Health Data from A Health Facility in Effia-Kwesimintsim Area of Sekondi-Takoradi, Ghana

Emmnauel Baah[1*] Frank B. K. Twenefour[1] Joseph K. A. Johnson[1]

1. Department of Mathematics and Statistics, Takoradi Polytechnic, P O Box 256, Takoradi, Ghana

* E-mail of the corresponding author: emmanuel.baah@tpoly.edu.gh

Abstract

Data gathered by our health institutions in the course of their work are an important source of information on disease patterns and their spatial distribution. Such information could be crucial in planning our public health policy. This paper reports on an analysis carried out on four kinds of data: Monthly Outpatient Morbidity Data, Statement of Outpatients, Statement of Inpatients and the Monthly Bed State Data of a health facility in Takoradi, Ghana. The data spans the six-year period of 1997 to 2002.

The objectives of the study were to: (i) Identify the leading diseases reported to the health facility, (ii) Identify the patterns, if any, in the diseases reported to the health facility, (iii) Explore the data with the view to finding any unsuspected characteristics.

The analysis of the data relied on descriptive statistics and time series analysis. Projected values of the number of cases of the diseases to be reported to the health facility, particularly that of malaria in the years following the period under consideration were explored.

The results from the analysis showed that Malaria is the leading of the 13 most common diseases in the catchment area with an upward trend in the number of cases reported to the health facility for the period under consideration. It was also found that the number of cases of Pneumonia and Hypertension were on the ascendancy in the period under consideration while tuberculosis is the most prevalent of the 'Six Child Killer Diseases' – Diphtheria, Poliomyelitis, Tetanus, Measles, Whooping Cough, and Tuberculosis – in the catchment area. Further, Females and the age group 0 – 4 years are the subgroups in the population of the catchment area that require the most medical attention. A probable causal factor in the dramatic rise in the number of malaria cases in 2001 was identified to be the construction of the Takoradi–Agona Nkwanta section of the Trans-West Africa highway which runs through the catchment area of the health facility.

Keywords: Time series, Forecast, Health statistics

1. Introduction

Good health is key to the development of the individual and the society in general. Authorities, be they local or national, therefore endeavour to provide services geared towards good health for its citizenry through various means including provision of healthcare facilities. An integrated health delivery system that employs all the approaches, including analysing the data gathered by our health institutions, in a bid to guarantee good health for the citizenry is central to total health delivery. Hence we present in this paper an analysis of data from a health facility in Takoradi, Ghana. The health facility, which is located on the right hand side of the road as one drives from Takoradi to Agona Nkwanta and close to Kwesimintsim town, serves people from about eight (8) communities scattered around it (Ghana, Ghana Health Service, 2003).

2. Review of Methods

This paper employs essentially two methods: Descriptive Statistics and Time Series Analysis. The descriptive statistics include measures used by people involved with health management though the interpretation of these statistics is done from the perspective of someone who has no background in health management.

2.1 *Descriptive Statistics*

The descriptive statistical tools used include ratios (some in the form of proportions and percentages), arithmetic mean and the geometric mean.

Geometric Mean

The geometric mean (*GM*) is useful in dealing with quantities that change over time, such as in finding the average growth rate of a quantity over a period of time. The geometric mean is preferred in such circumstances because it paints the true picture on the ground by giving a more conservative estimate than the arithmetic mean, which is unduly affected by extreme values (Keller and Warrack, 2003).

Health Descriptive Measures

The health facility like any other health centre under the Ghana Health Service calculates some Inpatient Statistics. Among other things they are used as Performance Indicators. They include the following (Ghana, Ghana Health Service, 2001):

a. *Available Bed Days (ABD)*

It is described as the Staffed Bed minus Temporary unavailable beds due to:

i. Shortage of staff.
ii. Minor redecorations.
iii. Outbreak of epidemics.

$$\text{ABD} = (\text{Bed Compliment}) \times (\text{Number Of Days})$$

b. *Patient Days (PD)*

It is the total number of beds in use or occupied in each ward per day for the period. It is also called the bed state. It is obtained by taking the cumulative totals of the beds state for each ward per day for the month, quarter or year.

c. *Vacant Available Bed Days (VABD)*

It is obtained by subtracting the total patient days from available bed days.

$$\text{VABD} = \text{ABD} - \text{PD}$$

d. *Average Length of Stay (Als)*

It is obtained by dividing the patient days by the total death and discharges for the period.

$$\text{ALS} = \frac{\text{PD}}{\text{Discharges} + \text{Death}}$$

e. *Average Daily Occupancy (ADO)*

It is obtained by dividing the total Bed Days by the number of days in the period.

$$ADO = \frac{\text{Total Patient Days}}{\text{Number Of Days}}$$

f. *Percentage Occupancy (PO)*

It is the total bed days divided by the product of the bed state and the number of days multiplied by 100.

$$PO = \frac{(\text{Patient Days}) \times 100}{(\text{Bed Compliment}) \times (\text{Number Of Days})}$$

g. *Turn Over Interval (TOI)*

It is the time that elapses between the discharge of a patient and the admission of another patient to occupy the same bed.

$$TOI = \frac{\text{ABD - PD}}{\text{Discharges} + \text{Death}}$$

h. *Turn Over Per Bed (TOPB)*

It is the total number of patients per bed per year.

$$TOPB = \frac{\text{Discharges} + \text{Death}}{\text{Bed Allocation}}$$

i. *Death Rate (DR)*

It is the ratio of deaths in the hospital during any given period of time to the total number of discharges and deaths during the period.

$$DR = \frac{(\text{Total Deaths}) \times 100}{\text{Discharges} + \text{Death}}$$

Kpedekpo and Arya (1981) had this to say on the importance of health statistics in general and the above statistics in particular:

> The average length of stay is usually regarded as a measure of the hospital's management efficiency. High occupancy rates are used as an argument for building extra capacity for the hospital. Information on out-patient attendances gives an estimate of illness and an index on the demand being made on the health services or indicates the workload of the staff. Rate of utilization of health facilities can be estimated from the population of a catchment area and the number of attendances. Naturally if there are great variations in this index within a country, this suggest inequalities in the distribution of health care resources (p. 59).

2.2 *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 examples 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

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 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 constant parameters – their true 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, ... 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):

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

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

MAD, MSD And MAPE

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

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

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

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).

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

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

$$\bar{R}^2 = 1 - \frac{S_{xy}^{\ 2}}{S_y^{\ 2}}$$

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

3. Results

An examination of the data shows that Malaria, Diseases Of The Oral Cavity, Upper Respiratory Tract Infection (URTI), Skin Diseases and Ulcers, Accidents, Diarrhoeal Diseases, Intestinal Worms Infestation, Gynaecological Disorders, Pregnancy and Related Complications, Chicken Pox, Hypertension, Mental Disorders and Pneumonia featured in the list of top ten causes of outpatient attendance at least once in the six years under consideration. Table 1 presents their ranks for the various years in the period under consideration.

Table 1: The most reported medical problems and their ranks

Disease	Rank					
	1997	1998	1999	2000	2001	2002
Malaria	1	1	1	1	1	1
Upper Respiratory Tract Infection	3	2	2	2	2	2
Skin Diseases And Ulcers	4	3	3	3	6	4
Accidents	5	6	4	4	7	10
Diarrhoeal Diseases	6	5	7	5	5	7
Intestinal Worms Infestation	7	7	8	9	3	5
Gynaecological Disorders	8	8	6	6	8	9
Diseases Of The Oral Cavity	2	4	5	-	4	3
Hypertension	-	9	10	7	9	6
Pregnancy And Related Complications	9	10	-	8	10	-
Mental Disorders	-	-	9	10	-	-
Pneumonia	-	-	-	-	-	8
Chicken Pox	10	-	-	-	-	-

The first seven of the diseases on the table were among the top ten diseases in all the six years under consideration.

It is clear that Malaria is the leading cause of attendance at the health facility. It featured in all the six years under consideration, taking the first position in each year and accounting for 37.65% of the total outpatient attendance (new cases) in 1997; the least for Malaria for all the years under consideration. The number of cases of Malaria rose steadily from 9,598 cases in 1997 to 13,513 cases in 2000, from where it shot up dramatically by 44.47 % to 19,523 cases in 2001 and then dropped in a similar fashion to 14,026 cases in 2002.

Infection of the upper respiratory tract (URTI) is the next leading cause of attendance at the health facility. It took the second position in the five of the six years under consideration, taking the third position in 1997.

The fact that the number of cases of URTI (being the second leading cause of attendance) could account for at most only 24.1% of the corresponding number of Malaria cases comparing the annual number of cases of the two diseases (excluding 1997) underscores the relative seriousness of Malaria.

Table 2 presents the composite figures of the inpatient statistics (health descriptive statistics) for the labour and gynaecology wards from 1997 to 2002.

Table 2: Composite figures of inpatient statistics

Inpatient statistic		Year					
		1997	1998	1999	2000	2001	2002
Death Rate	(DR)	0.8	1.0	0.6	0.1	0.1	0.4
Patient Days	(PD)	6925	5480	4662	3992	3755	3582
Average Daily Bed Occupancy	(ADO)	19.0	15.0	12.8	10.9	10.3	9.8
Occupancy %	(PO)	75.9	57.7	49.1	42.0	39.6	37.7
Average Length Of Stay	(ALS)	2.8	2.8	2.8	2.2	2.2	2.1
Turnover Per Bed	(TOPB)	100.1	76.4	64.1	70.1	66.3	65.5
Turnover Interval	(TOI)	.09	2.0	2.9	3.0	3.3	3.5

There is a general decreasing pattern in the values of PD, ADO, PO, ALS and TOPB while the values of DR appear to fluctuate. Also there is an increasing trend in the values of TOI.

To estimate the trend component in the number of malaria cases reported, the actual malaria series was deseasonalized (seasonal effect was removed from the actual series) by dividing the seasonal indices into the corresponding values of the actual series, that is $Y/S = T\ C\ I$. This is so because the presence of the seasonal effect tends to hide the trend in the data and explains why estimating the seasonal effect and deseasonalising the actual series is often the first step towards estimating the trend. The estimated trend was obtained by fitting a linear least squares regression equation, dealt with above, to the values of *TCI*. The model $\hat{y} = \tau s$, where τ and s are the estimates for the trend and seasonal components respectively, was assumed because the irregular and the cyclical components, if they exist, were thought to be insignificant.

It must be pointed out that there are several somewhat different approaches to decomposing time series. This is because $Y = T\ C\ S\ I$ has several identical forms which serve as the basis of the decomposition, but the basic principles are almost the same. The bottom line is developing a valid model that fits the data and yields forecasts with minimal error.

As pointed out above there is more than one approach to decomposing a time series to obtain estimates of the components and thereby be able to obtain forecasts. One way of decomposing a time series is as outlined above. Let us designate the model resulting from this approach as Model 1. Minitab uses a somewhat different approach:

Minitab first fits a trend line to the data, using least squares regression, then detrends the data by either dividing or subtracting out the trend component. Then it smooths the detrended data by using a centred moving average of length equal to the length of the seasonal cycle. If the seasonal cycle length is an even number, this actually requires a two-step moving average in order to synchronize the moving average correctly. Once the moving average is obtained, it is either divided into or subtracted from the detrended data to obtain what are often referred to as raw seasonals. Within each seasonal period, the median value of the raw seasonals is found. These medians make up the seasonal indices. The seasonal indices are in turn used to seasonally adjust the data (Minitab, 1996, Release 11.21).

We designate the model resulting from this approach as Model 2. These two models were fitted to the malaria series and were then compared.

Selecting and Assessing a Model

We compare the two models described above on the basis of forecast of the number of cases of malaria for 2003 using these models. The values of statistics needed to do the comparison are given in Table 3.

Table 3: Comparing Model 1 and Model 2 based on the accuracy measures

		Model 1	Model 2
Fit	*MAPE*	16.48	16.20
	MAD	179.42	177.30
	MSD	50809.78	52220.60
	$\overline{R}^2$ %	52.47	51.15
Forecast	*MAPE*	36.57	36.11
	MAD	386.11	386.95
	MSD	206009.08	190999.83

Considering the fits, it could be seen from Table 3 that the respective *MAPE* and *MAD* values of 16.20 and 177.30 for Model 2 are less than the corresponding values of 16.48 and 179.42 for Model 1, hence it appears Model 2 is the better of the two models; but the respective *MSD* and $\overline{R}^2$ values of 50809.78 and 52.47 for Model 1 are respectively less than and higher than the corresponding values of 52220.60 and 51.15 for Model 2, indicating that in terms of these two statistics Model 1 is the better of the two. However the differences in these statistics are not dramatic and since the values of the aforementioned four statistics are not all decisively in favour of any one of the models, one is inclined to conclude that in terms of the fits the two models are comparable.

Examining the forecast values of *MAPE*, *MAD* and *MSD* for the two models, the respective *MAPE* and *MSD* values of 36.11 and 190999.83 for Model 2 are slightly lower than the corresponding values of 36.57 and 206009.08 for Model 1; but the *MAD* value of 386.11 for Model 1 is lower than the corresponding value of 386.95 for model 2. Again the differences in the values of the statistics are small. However the fact that the values of *MAPE* and *MSD* for Model 2 is slightly lower than that of Model 1 makes it the favourable choice as a forecasting model.

We note that the forecasting accuracy of both models are inferior to the fitted accuracy, as the *MAPE*, *MAD* and *MSD* values of the fits are considerably lower than the corresponding values of the forecasts.

Turning to Model 2 as the model of choice, the mean error of – 338.89 indicates that there is some bias in the forecasted values, that is on the average a forecast value is 338.89 higher than the actual. We would want to establish whether or not this bias is statistically significant. Hence we test the hypothesis (alternative, H_1) $\mu_{e_t} > 0$ against the hypothesis (null, H_o) $\mu_{e_t} = 0$. Using the Wilcoxon Signed Rank Sum test with significance level $\alpha = 0.05$, the sum of the ranks of the residuals which were positive is given by $T^+ = 75$ and the sum of the ranks of the residuals which were negative is $T^- = 3$. Taking $T = T^-$ as the test statistic, the critical value $T_L = 17$. Thus $T < T_L$, hence we reject H_o and conclude that there is consistent bias, namely Model 2 is over forecasting by an average of 338.89.

4.0 Discussion and Conclusion

The study reveals that the 13 diseases, namely Malaria, Upper Respiratory Tract Infection (URTI), Skin Diseases and Ulcers, Accidents, Diarrhoeal Diseases, Intestinal Worms Infestation, Gynaecological Disorders, Diseases Of The Oral Cavity, Hypertension, Pregnancy And Related Complications, Mental Disorders, Pneumonia and Chicken Pox have been among the top ten causes of outpatient attendance at least once in the six years under consideration. The first seven of the aforementioned diseases appeared in the list of top ten causes of outpatient attendance in all the six years under consideration with malaria leading the pack, and accounting for 37.65% of the total outpatient attendance (new cases) in 1997; the least for malaria for all the years under consideration.

The seriousness of Malaria is underscored by the fact that the number of cases of URTI (being the second leading cause of outpatient attendance) could account for at most only 24.1.% of the corresponding number of Malaria cases comparing the annual number of cases of the two diseases (excluding 1997). However the dramatic rise in the number of cases of malaria in 2001 is believed to be unusual, possibly an agency might have provided the conducive environment for mosquitoes (the vector of malaria) to proliferate, leading to the dramatic rise in the number of cases of malaria. This line of thinking is all the more plausible when one considers the fact that a lot of pools of water – perfect condition for breeding of mosquitoes – resulted from the construction of the Takoradi – Agona Nkwanta section of the Trans-West Africa highway which runs through the catchment area between 2000 and 2002. Indeed much of the construction was undertaken in 2001. No wonder then, perhaps, that the number of cases of malaria dropped in a similar fashion when the road project was completed in 2002. The general trend in the data suggests that malaria cases will continue to rise in the years ahead.

The number of cases of Pneumonia in 2002 was more than two times the number of cases in all the preceding five years put together. An examination of the trend in the data suggests the number of cases will continue to rise after 2002.

There is an upward trend in the number of cases of Hypertension, one of the 'Diseases of Affluence'. Is it an indication that stress is on the increase in the catchment area or rather a change in the eating habits of the people of the catchment area? This is an issue that calls for further investigation.

Among the 'Six Child Killer Diseases' tuberculosis is the most prevalent. Though the number of cases of 46 is small compared to the leading diseases such as malaria, its debilitating and opportunistic nature – its one of the diseases that capitalise on the weakened immune system of HIV/AIDS patients to kill them – makes the need to control and eradicate it urgent.

We also conclude from the analysis of the Monthly Outpatient Morbidity Data for the six years under consideration that the Polyclinic has an average annual utilization of 26.6% (the geometric mean of the annual utilization rates) for the period under consideration.

Apart from 1997, the proportion of patients from the age groups $0-4$ years and $5-14$ years were respectively the highest and the smallest in any given year within the six-year period under consideration. Also the proportion of female patients was higher than that of the males in all the years under consideration. It follows that the age group $0-4$ years and females are the subgroups in the catchment area that require the most medical attention while the age group $5-14$ years require the least attention.

The patterns in the values of the inpatient statistics appear to raise a number of questions. The increasing trend in the values of TOI means that the interval between when a patient is discharged and when another is admitted to occupy the same bed is increasing. Could this mean that the general health situation of women in the catchment area is improving or the people in the catchment area are turning to other health facilities, for which reason there is a let up in the demand rate (in terms of TOI)? Are the programmes on the local electronic media that focus on women impacting positively on female health in the catchment area? This also calls for further investigation.

Granted that patients who are discharged have really been cured of the medical problems they brought to the health facility, the decreasing trend in the values of ALS appear to be an indication of the improving management efficiency at the health facility. The increasing trend in the values of TOI and the decreasing trend in the values of ALS seem to be in line with the decreasing trend in the values of TOPB. Obviously if the length of stay is becoming shorter and it is taking longer to admit a patient to occupy the same bed a previous one has left, then the demand on the beds is expected to decrease.

The decreasing trend in the values of PO and ADO seem to be consistent with the decrease in ALS and appear to suggest that the need for extra capacity is not as pressing as it used to be or the need for extra capacity is decreasing. The fluctuating values of DR suggest that the death rate is not stable.

The model of choice, Model 2, was found to be over-forecasting by an average of 338.89. It is believed that the trend might have been unduly influenced by the unusually high number of cases of malaria in 2001 leading to the forecasts for 2003 being higher than the actual values, and that it is only a matter of time that the differences between the forecasted values and the actual values will take on values which will render the mean error not significantly different from zero, all things being equal, and the model will prove useful. However over-forecasting, in the context of the data used is not as serious as under-forecasting which would have meant that the Polyclinic could run out of resources.

In picking a time series model to be used to get projected values of the number of cases of the diseases in the immediate future, the health facility could employ the analysis described above to select any one of the two models, whether both models are based on the same kind of trend, like linear-linear, quadratic-quadratic et cetera but differ in the approach used in the decomposition; or the two models are based on different kinds of trends, like linear-quadratic, linear-exponential et cetera but are based on the same approach of decomposition; or both are based on different trends and different procedures of decomposition.

We conclude by pointing out that study shows that timely analysis of data gathered by health institutions could be helpful in revealing the trend in the occurrence of diseases and their link to what is happening in the environment, which could provide the information needed to properly plan public health policy.

References

Bowerman, B. L. and O'Connel, R. T. (1997), *Applied Statistics: Improving Business Processes.* Bonston, USA: Irwin/Mcgraw-Hill.
DeLurgio, S. A. (1998), *Forecasting Principles and Applications*. 1st ed. Bonston, USA: Irwin/Mcgraw-Hill.
Freund, J. E. (2001), *Mathematical Statistics*. 5thed. New Delhi: Prentice-Hall.
Ghana. Kwesimintsim Polyclinic, Ghana Health Service (2001), *Annual Report.* Takoradi, Ghana: Kwesimintsim Polyclinic.
Ghana. Kwesimintsim Polyclinic, Ghana Health Service. (2003), *Half Year Report.* Takoradi, Ghana: Kwesimintsim Polyclinic.
Ghana. Ghana Statistical Service. (2000), *Population and Housing Census*. Accra, Ghana: Ghana Statistical Service.
Harper, W. M. (1991), *Statistics*. 6th ed. London, UK: Longman Group Ltd.
Keller, G and Warrack, B. (2003), *Statistics for Management and Economics*. 6thed. Madrid, Spain: Thompson, Brooks/Cole.
Kpedekpo, G. M. K and Arya, P. L. (1981), *Social and Economic Statistics for Africa*. 1st ed. London, UK: George Allen & Unwin Ltd.
Spiegel, M. R. (1992), *Statistics*. 2nd ed. London, UK: McGraw-Hill International (UK) Ltd.
USA, Minitab Inc. (1996), *How Minitab Does Decomposition*. Minitab Release 11.21. PA, USA.