Time Series Analysis of Children born with HIV at Bwaila Antenatal Clinic, Lilongwe

Applied Statistics

Dissertation

Joachim C. Kapalamula

Faculty of Science

The Catholic University of Malawi

July 2021

The Catholic University of Malawi

Faculty of Science

Time Series Analysis of Children born with HIV at Bwaila Antenatal Clinic, Lilongwe

Applied Statistics

Dissertation

Joachim C.Kapalamula

BScSTAT/NE/2017/002

Submitted to the department of Mathematical Sciences in partial fulfilment of the requirements for the award of a bachelor of Science degree in Applied Statistics

July 2021

Declaration

I the undersigned hereby declare that this thesis constitutes my original work and has not been submitted at the Catholic University of Malawi or elsewhere for academic purposes. The references made in this dissertation have been acknowledged accordingly.

Joachim Chikuni Kapalamula

Full Legal Name

Signature ..... **Date .....**

Certificate of Approval

The undersigned certify that this thesis represents the student's own work and effort and has been submitted with our approval.

Mr H.Banda
Supervisor

Signature ..... **Date .....**

Mrs M.Chitete
Head of Department
Mathematical Sciences

Signature ..... **Date .....**

Mr I.Chikopa
Dean
Faculty of Science

Signature ..... **Date .....**

Dedication

To my Mom, my late Aunt and to Daddy.

Acknowledgements

I owe it all to God for being with me throughout my studies, thank you Lord.

Special appreciation to my parents respectively, for their profound support in all aspects you made it all conducive, I shall remain grateful.

I’m so humbled for all support from my supervisor and all stuff in the Mathematical Sciences Department throughout my academic journey, my classmates and friends and the entire Statistics family on Campus.

As well as the Directorate of Health and Social Services at Bwaila Antenatal Clinic for their data endorsement, may God abundantly bless you all.

Abstract

This study was an attempt to apply Time series analysis on Children born with HIV using SAS software version 9.4. The ARIMA methodology developed by Box and Jenkins was used in this paper. Time Series refers to an ordered sequence of values of a variable at equally spaced time intervals.

Time series occur frequently when looking at the data applications. The analysis was carried out using time series data on the Children born with HIV at Bwaila Antenatal Clinic in Lilongwe District covering a period from 2013 to 2020 which was collected from the office of the Healthy Information management systems in Lilongwe at the District Healthy Office. The results from Empirical studies outlined a decrease in birth of infants through MTCT of HIV, the analysis of quarterly HIV exposed births of Children data was used to identify the pattern.

The non-seasonal pattern for HIV exposed births was at highest in 2016 second quarter followed by second quarter of 2020 and there were no births in 2013 second quarter such that it recorded zero. It was further inferred that the forecast for the Children born with HIV, had approximately the same values for the next four years, thus 6.3 in the first quarter 2021 with confidence limits (-3.2 to 15.8) while the preceding quarters had 6 births forecast with (-3.3 to 15.7) confidence limits until the last quarter of 2025. The results in this study suggested that the number of HIV exposed infant births had reduced through the strategies enhanced but there won't be a complete elimination towards MTCT due to other personal and service delivery challenges.

The forecasting had shown a constant horizontal trend in births of infants through Mother to Child Transmission hence this provides sufficient evidence that HIV births are on reduction locally considering Bwaila Antenatal Clinic births but, we are not close towards elimination.

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

ARIMA	Auto-regressive Integrated Moving Average
ADF	Augmented Dickey-Fuller Unit root test
ACF	Auto - Correlation Function
AIC	Akaik Information Criteria
ART	Antiretroviral Treatment
CD4	Cell Count
HIV	Human Immune Deficiency Virus
HIMS	Healthy Information Management Systems
MTCT	Mother to child transmission
MCH	Maternal Child Health
NAC	National AIDS Commission
NVP	Neverapine
OPD	Outpatient Department
PMTCT	Prevention of Mother to child transmission
PACF	Partial Correlation Function
SAS	Statistical Analysis Software
SBC	Schewatz Bayesian Information Criteria
sdNVP	Single-dose Nevirapine
UNAIDS	United Nations Acquired Immune Deficiency Syndrome
UNICEF	United Nations International Children's Emergency Fund
VCT	Voluntary Canceling and Testing
WHO	World Health Organization

List of Symbols

$L(lags)$ ..... *The time period between two observations*

q *or* p ..... *Number of lags*

Y_{1-t} ..... *Observation at time* $t-1$

t ..... *Time Component*

a_{1-t} *or* E_{1-t} ..... *Random errors*

Θ ..... *Theta*

Φ ..... *Phi*

(p,d,q) ..... *Auto − regressive, Differencing and Moving Average Factors respectively*

Chapter 1

Introduction

1.1 Background Information

Babies are born with HIV from their mothers vertically through a process called MTCT. Mother-to-child transmission of HIV (MTCT) is defined as the transmission of HIV from a mother to her child during pregnancy, labor, delivery or during breast feeding, as stated by an online Journal HIV AIDS (Auckl) version v.11; 2019. In a report, UNAIDS Geneva, Switzerland (1999) states that, in the absence of preventive measures, the risk of a baby acquiring the virus from an infected mother ranges from 15 to 25 percent in industrialized countries, and 25 to 35 percent in developing countries. The difference is due largely to feeding practices, breastfeeding is more common and usually practiced for a longer period in developing countries than in the industrialized world, UNAIDS/99.40E (English original, August 1999).

The prevention strategies are set to deal with the MTCT of HIV. A study by the World Health Organization Regional Office for Africa Brazzaville 2014 on Implementation of Option B+ for Prevention of Mother-To-Child Transmission of HIV: The Malawi experience, states that there has been an optimal effectiveness in eliminating MTCT over the years, thus PMTCT. The strategies on PMTCT in Malawi started in 2001, initially were being piloted at Embangweni Mission Hospital in Chiradzulu District and Thyolo District Hospital. In Lilongwe district, implementation of PMTCT services started in April 2002 with 4 health facilities.

According to Malawi Medical Journal, 2017 Dec; 29(4), the coverage of PMTCT services had expanded rapidly and initiatives to improve delivery of PMTCT services have been equally impressive. Despite the effectiveness of the strategies, there is still a challenge in that a fraction of infants being born with HIV still exists.

1.1.1 Malawi National AIDS Policy

The government of Malawi through a report on National HIV/AIDS policy: A call for renewed action in October 2003, stated that, the desire of HIV-infected couples to have a child must be

balanced with the possibility of having an HIV-infected baby who has a high risk of dying in early childhood, after suffering extended periods of illness (Morbidity). In addition, the report states; that the death of a parent, especially the mother, drastically reduces the baby's chances of survival, regardless of the baby's HIV serostatus. It is important, therefore, that interventions addresses treatment for parents, in addition to PMTCT, so as to minimize orphan hood and improve the chances of child survival.

To ensure the statements above, the government stipulated policy statements to terminate MTCT in 2003 as of then, PTMCT was just officially being launched through NAC. The policy encouraged on;

- Promoting VCT for couples planning to have a Child, and early attendance at Antenatal Clinics.
- Ensuring that HIV testing is routinely offered to all pregnant women attending Antenatal Clinics unless they specifically choose to decline.
- Ensuring the availability of quality infrastructure, skilled staff and supplies for Maternal and Child Health (MCH) care, and proper management of MCH services to increase women's access to PMTCT interventions.
- Providing access to affordable Antiretroviral Treatment (ART) to prevent HIV transmission from Mother to Child, and ensuring that PMTCT programs also provide treatment, care and support for both parents.
- Providing an enabling environment for women to participate in PMTCT or other preventive care or support programs without the consent of their husbands, sexual partners or family.
- Ensuring baby-friendly hospital initiatives to support HIV-positive lactating mothers who choose to exclusively breastfeed for six months.

However, The Malawi Experience Report on Implementation of Option B+ on MTCT pointed out that, as the national PMTCT program was officially launched in 2003. Single-dose Nevirapine was used for the period 2001 to 2008. Malawi introduced AZT combination prophylaxis in 2008 with AZT/3TC being initiated at 28 weeks, followed by sdNVP during labor and AZT/3TC tail for a week. Infants received NVP syrup for a week or more depending on the duration of exposure of the mother to ARVs. At the same time, sDNVP for mothers and infants was phased out in Malawi.

The report further states, Malawi convened a stakeholder's symposium in January 2010 to examine a feasible option from the new WHO recommendations. Following a recommendation from the stakeholder's symposium, Malawi decided to put all HIV positive pregnant and

breastfeeding women on ART for life irrespective of the clinical stage or CD4 cell count. This approach is now known as Option B+ which is being implemented since July 2011. In April 2012, the World Health Organization published a programmatic update that recognized the potential advantages and the need to take into consideration the new **Option B+** approach within a public health perspective in resource-limited settings?

1.1.2 Significance of the Option B+

The importance of the option B+ approach is to provide a simplification of PMTCT and ART treatment regimens and service delivery; strengthening of linkages between reproductive health and ART programs at all service delivery levels at the Antenatal Clinics; towards protection against MTCT in future pregnancies and between discordant couples; and avoiding stopping and re-starting ARVs, in other words it involves the use of ART for life.

Option B+ also enables safe breastfeeding while avoiding the need for prolonged infant HIV prophylaxis which is challenging. However, on the other hand it also enables women become pregnant again after a breastfeeding period of 23 months, given the high rates of fertility in the country and the short time intervals between pregnancies.

1.2 Statement of the problem

However it is believed that many HIV-positive births have been prevented through these prevention strategies implemented. This statement is also witnessed by the research conducted by UNICEF in February 2016, a case study on joint community-facility review of PMTCT dashboards in Malawi, which states that;

> Malawi has made good progress towards the goal of eliminating mother-to-child transmission of HIV. In just four years, the country has halved the number of new infections among children from 21,000 in 2010 to 10,000 in 2014.

In the sustainable development goals of Malawi, goal number 3 states to ensure health lives and promote well-being for all at all ages. More precisely, target indicator 3.1.2 ensures that proportion of births should be attended by skilled health personnel. This is why we have the projects such as the Prevention of Mother to Child Transmission (PMTCT) to ensure that Mother to Child transmission is reduced.

Although reports on effectiveness of PMTCT are optimal about MTCT births, the data on Mother to Child Transmission has never been traced through patterns of time series analysis locally in the country to give an assessment and a picture on how MTCT has been since implementation of these PMTCT strategies. On the other hand, these reports have aimed at proving the effectiveness of the strategies on PMTCT of a given particular period of time without giving

its prognosis in the long-run. However in addition to that, the reports were specific on just some selected Antenatal Care Centers while being categorized by short durations.

The problem of not using time series in issues like these is that the practioners, stakeholders and the government fails to see through the whole picture behind an interventional strategy over time. The analyses might have been there to give a national, regional or an international and a Hospital collection insight but if it cannot tell a local significance pictorially then it is a problem of deficiency in evaluation. There are no existing reports on Time series Analysis in Malawi on Children born with HIV and hence this has never been addressed locally in as far as pediatrics research is concerned in the country.

Hence the study to be conducted at Bwaila Antenatal Clinic in Lilongwe, considering appropriate statistical procedures and applications on time series.

1.3 Purpose of the Study

The study aimed at focusing on the basic and accurate concept of analyzing the time series data locally of the Children born with HIV at Bwaila Antenatal Clinic, using time series models.

1.4 Objectives of the Study

The following were the objectives for the study.

1.4.1 Main Objective of the Study

The study performed a Time Series Analysis of Children born with HIV at Bwaira Antenatal Clinic in Lilongwe District.

1.4.2 Specific Objectives of the Study

- To identify the basic Time Series approach technique upon analyzing the time series data of the Children at Bwaila District Hospital.
- To examine the trend/pattern in the time series data.
- To determine an appropriate and an accurate Forecasting model for the time series data at Bwaila.

1.4.3 Research Questions

The following were the likely research questions for the study,

- What was the basic Time Series approach techniques used upon analyzing the time series data of the Children at Bwaila District Hospital?
- What was the trend/pattern in the time series data?
- What was the appropriate and accurate model for Forecasting the Time Series data at Bwaila?

1.5 Significance of the study

In general the study will give a picture on the changes in the number of children born with HIV through Mother to Child Transmission with time at the Hospital.

On the other hand the study had a scholarly benefit, however it is hoped that the findings will also be important to the government, such that measures will be taken to improve the health of its future tax payers and reduce mortality and morbidity in the process. The study will help policy makers and Social organizations to find an intervention or change their intervention strategies in the field of prevention of Mother to Child Transmission through the information obtained from the study.

The District Healthy Office at the hospital will be able to find out the stage where it stands in fighting against MTCT over time and hence update its interventions strategies, the community on the other hand will be able to either trust the services offered at the hospital or otherwise.

1.6 Limitations of the Study

The study considered to use data ranging from 2002 - 2020 therefore electronic data for a considerable time period on MTCT was not available hence, the study used quarterly data that was available from 2013 to 2020.

Besides, the protocols at the hospital does not allow diagnosis of fresh babies for HIV status therefore the data is collected in the first follow-up of Antenatal visit after child birth in cohorts, where some couples do not even attend hence, a baby can be born negative but may contract HIV through interactions with the parents while at home which complicates the chemistry of MTCT.

Despite data restrictions due to legal and policy reasons were another issue, the study had an Academic platform and hence the data were granted from the HIMS office at Bwaila.

Chapter 2

Review of Related Literature

2.1 Theoretical Review

The study used the ARIMA modeling approach, a theory developed by George Box and Gwilym Jenkins in 1976 for time series data (North Dakota State University). ARIMA assumes a more accurate forecast since it is a combination of time series methods. On the other hand ARIMA is a widely used statistical method for forecasting. The main issue in ARIMA model is that it enables choosing lags and then decide which combination would be best for the time series data, Pollock, 1992. The time series data is plotted after which the patterns are observed in the data, since the series has to be stationary differencing is performed in order to remove the non-stationarity components such as the trend, random walks and the outliers.

Univariate ARIMA is a technique of forecasting, which refers just to its own series while doing the prediction (Morisson, n.d.). The model reviews time sequence graphs and explains how inspection of these plots enables the analyst to examine the series for outliers, missing data, and stationarity. It expounds graphical examination of the effect of smoothing, missing data replacement, and/or transformations to stationarity. Correlogram review also permits the analyst to employ other basic analytical techniques, allowing identification of the type of series under consideration, Robert Y.A. (1999).

2.1.1 Box-Jenkins ARIMA Modeling Approach

The Box-Jenkins model uses iterative three-stage modeling approach according to Barbara G.T and Linda S.F (2013) which includes;

1. Model identification and model selection: This makes sure that the variables are stationary, identifying seasonality in the dependent series (seasonally differencing it if necessary), and using plots of the autocorrelation and partial autocorrelation functions of the

dependent time series to decide which (if any) autoregressive or moving average component should be used in the model.

2. Parameter estimation using computation algorithms to arrive at coefficients which best fit the selected ARIMA model. The most common methods use maximum likelihood estimation or non -linear least-squares estimation.

3. Model checking by testing whether the estimated model conforms to the specifications of a stationary univariate process. In particular, the residuals should be independent of each other and constant in mean and variance over time (plotting the mean and variance of residuals over time and performing a Ljung-Box test or plotting autocorrelation and partial auto-correlation of the residuals are helpful to identify misspecification). If the estimation is inadequate, we have to return to step one and attempt to build a better model.

2.1.2 Configuring the ARIMA Model

ARIMA has two parts which are Autoregressive Models and Moving Average models.

2.1.2.1 The Autoregressive Models (AR)

This involves that observation of the series at Y_t can be explained by some function of its previous observation at Y_{t-1} adding the error variable, that is E_t (Morisson, n.d.). Thus, this means that it is possible to forecast the Y_t value with having Y_{t-1} and all the other necessary constants and figures, that are derived from time series. (Morisson, n.d.), Douglas C.M et al,(2008).

The AR is given by

$$Y(t) = \Phi(1) * y(t-1) + E(t) \tag{2.1}$$

Where,

Y_t = Time series under investigation, $\Phi(1)$ = The autoregressive parameter of order 1, y_{t-1} = The time series lagged 1 period and E_t = The error term of the model

2.1.2.2 Moving Average (MA)

The second part of ARIMA model is the Moving averages.

The difference between the autoregressive model and the moving average model is that the moving average model will put more focus on error constant of the previous observations or also called previous lags of the series; as considered by the equation shown below:

$$Y_t = -\Theta 1 * E_{t-1} + E_t \tag{2.2}$$

Where;

$\Theta 1$ is an MA of order 1 and it is multiplied with the error term of lag one. Therefore, according to moving averages model Y_t, which is the forecast future value, is always dependent on the error term.

2.1.2.3 ARIMA Model

The ARIMA Model mixes both equations together to conduct a more accurate forecast.

The auto-regressive, integrated, moving average model of a time series will be defined by three terms (p,d,q). Identification of a time series is the process of finding integer, usually very small (e.g., 0, 1, or 2), values of p, d, and q that model the patterns in the data. When the value is 0, the element will not be needed in the model. The middle element, d, is investigated before p and q. The goal is to determine if the process is stationary and, if not, to make it stationary before determining the values of p and q, (Morisson n.d.) , Barbara G.T, (2008) and Box G., and Jenkins G., (1976).

The model follows the equation below.

$$Y_t = \Phi_1 y_{t-1} + ... + \Phi_p y_{t-p} + a_t - \Theta_1 a_{t-1} - ... - \Theta_q a_{t-q} \tag{2.3}$$

using the back shift operator $By_t = y_{t-1}$
equation (3) reduces to;

$$\Theta_q(B) = (1 - \Theta_1 B - ... - \Theta_q B^q \tag{2.4}$$

where the a_t are random errors similar to E_t

2.1.3 Parameters of the Model

The model considers the following parameters;

- Auto - regressive (p), the number of lag observations included in the model, also called the lag order.
- Integrated ($trend$ (d)), the number of times that the raw observations are differenced, also called the degree of differencing.
- Moving average (q), the size of the moving average window, also called the order of moving average.

- Y_t ; an observation at time (t)
- Lag (L) ; the time periods between two observations.
- Θ ; the Moving average order
- Φ; the autoregressive order
- B; the back shift operator.

2.1.4 Assumptions

The Time Series ARIMA model assumes that;

- The series should be stationary, if not apply differencing or other models should be used.
- Normality of Distributions of Residuals
- Homogeneity of Variance and Zero Mean of Residuals
- Independence of Residuals
- Absence of Outliers

2.2 Empirical Review

Several studies have depicted a relationship existing between MTCT and PMTCT, most research studies have depicted that, MTCT births depends on PMTCT. For example, M van Lettow et al in a research on Prevention of mother-to-child transmission of HIV: a cross-sectional study in Malawi concluded that;

> The risk of MTCT is increased if any of the main steps in the program were missed.

As most pregnant positive women adheres to ART and attends to Antenatal Clinics at the same time, they prevent the spread of HIV to their infants through transmission at delivery and thereafter. This research however, visualized MTCT data at Bwaila, locally, through patterns of time series, from 2013 through 2020 in order to observe the changes in the patterns and forecasting despite the programs.

Recent researches have focused out on prevention of Mother to child transmission of HIV and evaluations following through different ways of prevention strategies opted such as Antenatal HIV testing, offering HIV-infected pregnant women life-long Antiretroviral Therapy (Option B+), assuring exclusive breast feeding for at most 2 years for HIV infected mothers, infant Antiretroviral prophylaxis and early infant HIV testing. While it has been likewise, most studies

have shown effectiveness of PMTCT at national, regional and international level perspective towards mother to child transmission births. On the other hand the time coverage of most of these analyses just gave a short time interval consideration. Few studies had focused on MTCT while observing PMTCT measures, such that they made observations over a short period of time while on the other hand considered a grouped set of some Antenatal Clinics to obtain an insight altogether.

2.2.1 Traditional Studies Reports

A study by M.A.Sinunu, E.J. Schouten, et al on Evaluating the Impact of Prevention of Mother-to-Child Transmission of HIV in Malawi through Immunization Clinic-Based Surveillance, with the aim of describing a surveillance approach to obtain population-based estimates of the vertical transmission rate VTR of infants less than 3 months of age in Malawi immediately after the adoption of Option B+. A sample of caregivers and infants less than 3months from 53 randomly chosen immunization clinics in 4 districts were enrolled and blood samples were tested to determine infant infection status and VTR. Caregivers were surveyed about maternal receipt of PMTCT services. The results were that Of the 5,068 blood samples, 764 were positive indicating 15.1 percent of mothers were HIV-infected and passed antibodies to their infant. Sixty-five of the positive samples tested positive, indicating a vertical transmission rate of 8.5 percent. Survey data indicates 64.8 percent of HIV-infected mothers and 46.9 percent of HIV-exposed infants received some form of Antiretroviral prophylaxis. Concluded that the observed VTR was lower than expected given earlier modeled estimates, suggesting that Malawi's PMTCT program has been successful at averting perinatal HIV transmission.

Consequently, in another study by M Van Lettow, M.Landes et al, on prevention of Mother to Child Transmission of HIV: a cross-sectional study in Malawi, aimed to estimate the use and outcomes of Malawian program for prevention of MTCT of HIV. In other words determining the effectiveness of option B+ strategy in a nationally representative sample mother-infant pairs enrolled at 4 weeks to 26 weeks postpartum. The sampling flame included all 579 healthy facilities that provided PMTCT services in Malawi in 2012 to 2013. Out of 34637 mothers and caregivers interviewed and tested for HIV 1.9 percent were excluded because of non-eligibility and 0.7 percent because of inconsistent answers on PMTCT. The rest containing 33744 mothers who were attending care with infants between 4-26 weeks old were included.

Upon the study the estimated uptake of Antenatal testing was 97.8 percent while maternal Antiretroviral Therapy was 96.3 percent; infant prophylaxis was 92.3 percent; and infant HIV testing was 53.2 percent. Estimated ratios of MTCT were 4.7 percent overall and 7.7 percent for the pairs that had missed maternal Antiretroviral Therapy, 10.7 percent for missing both maternal Antiretroviral Therapy and infant prophylaxis and 11.4 percent for missing maternal antiretroviral therapy, infant prophylaxis and infant testing. Women younger than 19 years

were more likely to have missed HIV testing. The study concluded that most women used the Malawian program for the prevention of MTCT. The risk of MTCT is increases if any of the main steps in the programs were missed. It further reads that;

> The results add to the evidence indicating that, within Malawi, the scale-up of the implementation of the Option B+ strategy has provided widely decentralized and equitable coverage of PMTCT services. The elimination of pediatric HIV infections will probably depend on the full use of PMTCT services.

2.2.2 International Study Reports

Additionally, a UNAIDS report in 2016 on Malawi estimates concerning coverage versus transmission of HIV, indicated an increase coverage of Antiretroviral medicines which translated into a decrease in rates of HIV transmission between the years 2009-2015 as shown by figure 2.3.

The other, figures 2.1 and 2.3 respectively presents summaries of child infections due to gaps in prevention of vertical transmission in Eastern and southern Africa 2017 and 2019 as concluded by UNAIDS epidemiological estimates 2020, (`https://aidsinfo.unaids.org/`). The greater the gap in prevention, the higher the child infections on African regional level.

Unite for Children Unite Against AIDS: Countdown to zero, Elimination of new infections among children by 2015 and keeping their mothers alive; UNAIDS, unpublished HIV estimates 2012 global report indicates that "Globally, an estimated 330,000 children were newly infected with HIV in 2011 down 24 percent from the 430,000 new infections in 2009. New pediatric HIV infections rose consistently until peaking at 560,000 in 2002 and 2003. Despite the significant progress that has been achieved, much more progress is needed in order to achieve that Global Plan target of a 90 percent reduction in the number of new HIV infections in children by 2015". See figure 3.

These studies altogether and many other more researches, have conveyed a similar goal in line with MTCT. Despite M.A.Sinunu's et el report, evaluated the impact of PMTCT in Antenatal Clinics based on the four districts of Malawi and further that, M Van Lettow's Cross-sectional study focused on the PMTCT representing the country at large. On the other hand UNAIDS report in 2020 just targeted on the Eastern and southern parts of African continents, in additional, there were also global reports on the issue at hand. However, the time coverage and intervals of these studies were insignificant in evaluating the strategies towards MTCT births. Despite that and many other factors, the studies have concluded that there has been an observable success in prevention of MTCT towards reduction in infants born with HIV from their positive mothers.

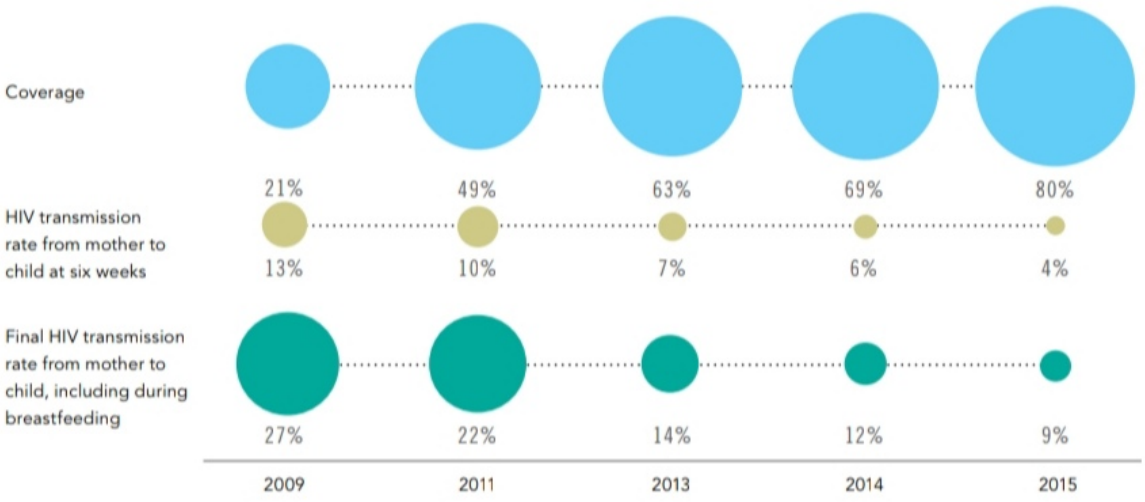
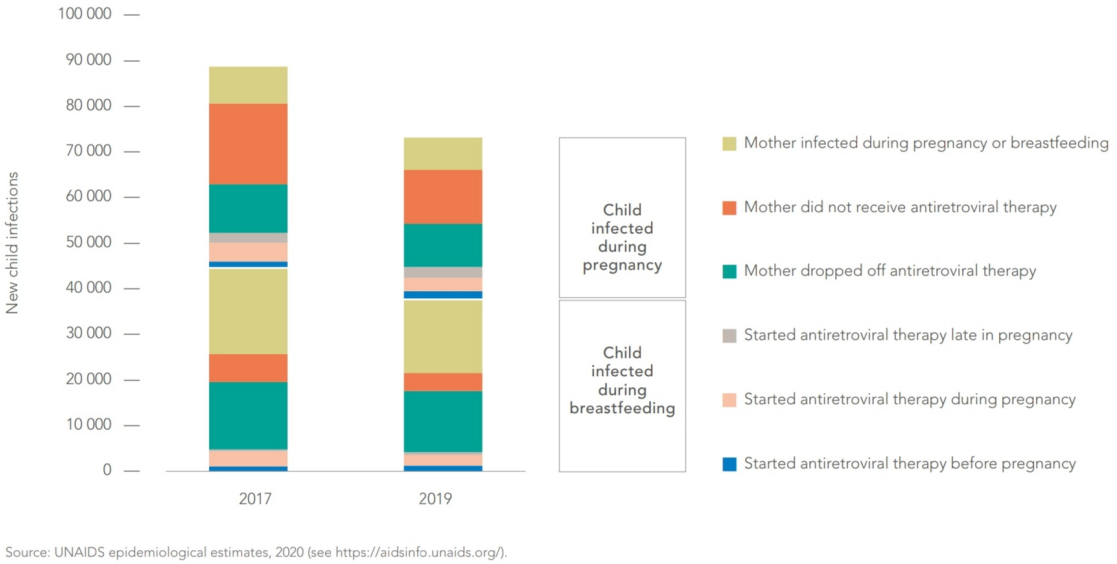


Figure 2.1: ART coverage Versus HIV transmission



Source: UNAIDS epidemiological estimates, 2020 (see https://aidsinfo.unaids.org/).

Figure 2.2: East and Southern African child infections due to gaps PMTCT

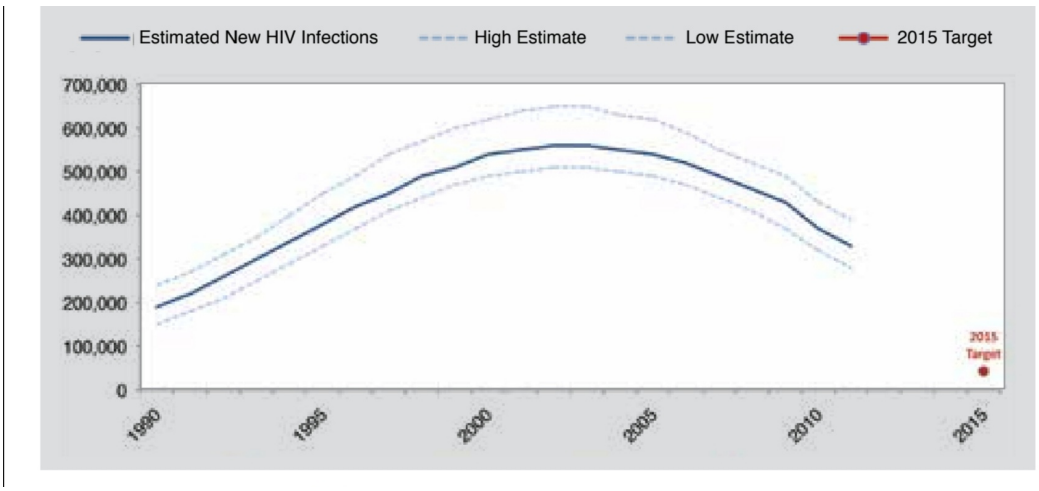


Figure 2.3: Estimated HIV infections globally ages 0-14 of 1990 - 2011, a 2015 target

In the changing world and considering various methods in research, it is not enough to evaluate the levels of PMTCT just by observing prevalence on a particular period of time or a year, mostly considering a short time coverage to see through mother to child transmission. It is also not enough to obtain evaluations centered on national level and other considerations before looking at the evaluations on the ground locally. The research findings are relevant but they cannot help to see on how to go about in the future and in addition to that, updating the relevant intervention strategies in the fight against HIV vertical transmissions among infants.

It is therefore the idea of this study to run through a time series analysis on children born with HIV at Bwaila Antenatal Clinic, considering that it is a local clinic in the district of Lilongwe. The analysis considered a time flame from 2013 to provide a summary picturing how the changes have been over the time. The study considered only the data of infant's HIV births at the Clinic because empirically it is not easy to obtain a quantitative data on PMTCT. Therefore using time series, we will assume an ARIMA model, plotting the dataset and analyzing through the patterns and finally performing a forecast based on the data.

Chapter 3

Research Methodology

3.1 Study Area

Bwaila Antenatal Clinic is found in the central region of Malawi in Lilongwe district, it is open 24hrs and covers 253km, travelled within 4hrs 47mins from Blantyre. It is 200 fit by 100M wide, (googlemaps, jul-15-2021, 7pm).

It is the main referral for hospitals in the capital city of Malawi with approximately 200 sick beds and receives an approximation of 176 clients per day on average, (Bwaila OPD). Bwaila hospital is considered to be one of the early selected few Antenatal Cares among the 585 sites as depicted in 2012 that were involved in undertaking Mother to Child Transmission services in the country, as stated by MalawiMedicalJournal,2017Dec;29(4). Among them were also Chiladzulu District Hospital and many others.

Therefore Bwaila possesses a great credibility of being involved in the study, and considering it being one of the earliest hospitals in the country it holds necessary representativeness.

3.2 Study Design

The study was Cross-sectional and quantitative in nature, where Secondary time series data from Bwaila Antenatal Clinic on Children born with HIV was used. A cross-sectional study involves looking at data from a population at one specific point in time, where the subjects are selected based on particular variable of interest, Bacon-Shone.J (2013),

3.3 Sampling Techniques

3.3.1 Sampling Frame

The study included all children born with HIV at Bwaila Antenatal Clinic from 2013 to 2020.

3.4 Data Collection

The study used secondary data from Bwaila hospital on Children born with HIV in the department of Health Information Management systems Database (HIMS) through the District Healthy Office of the hospital from 2013 quarterly to 2020.

The data at the facility is kept in cohorts that reports after birth for follow-up visits in the hospital, a cohort is a group of individuals having a statistical factor in common in a demographic study, Kang.S (2007).

3.4.1 Statistical Analysis

The data was formatted, cleaned and checked for outliers and missing values. Statistical Analysis Software version 9.4 (SAS) was used for data analysis set at 95% level of significance.

3.4.2 Statistical tests

- Dickey-Fuller test statistic; Tests for stationarity of the Time series data. To determine whether the series is stationary or not we considered the graph of ACF. If a graph of ACF cuts of fairly quickly or dies down fairly quickly, then the time series value should be considered stationary. With the ADF test we observe the Tau statistic with the H_0 and H_a that the series is Non-stationary and it is stationary respectively.

- Goodness of Fit; the model was estimated Using the maximum likelihood procedure to determine the parsimonious model, the coefficient less than 1 or -1 was detected to be a significant coefficient. On the other hand the Model with the lowest AIC (akaik information criteria) or SBC (Schewatz Bayesian Information Criteria) after checking was considered the best model for the forecast.

- Tests for Normality and Residuals; the study examined the normality of the residuals as the indicators of the properly fitted model, by considering the histograms distributions and the Q-Q plots of the residuals. Thus if the histogram are approximately normal, the model is a good fit.

3.4.3 Ethical Consideration

Watts, D. (1997) states that, from an ethics perspective, if the researcher is to value and respect the contributions made by participants, funding bodies and others supportive of the research effort, it is incumbent on the researcher to report and disseminate the findings of the particular study in the most effective ways available to the researcher.

Watts, D. (1997) further inquiries that, in reporting the study results, the ethical issues include continued protection of the rights of, and honoring promises made to, participants (for example, confidentiality, protection of privacy, anonymity), reporting findings truthfully, accurately and completely, citing appropriately the work of others and ensuring the authorship credits and acknowledgements are stated accurately.

To do otherwise indicates lack of respect for the various actors in the research process.

> It is also wasteful of valuable resources, including those of future researchers who might have gained from the signposting of 'blind alleys' and from insights into the findings, strengths and weaknesses of the unreported study.

The objectives of this study were therefore carefully examined at both the HIMS representatives and the Hospital's Research Board.

An approval letter was obtained from The Catholic University of Malawi and data permission letter was obtained from the Director of Healthy and Social Services at the hospital. The participants information were kept anonymous in the sense that not even a name of a subject was disclosed to the researcher as only data with respect to time was only appropriate for the study.

Chapter 4

Results and Discussion of the Study

In this chapter, the results of the time series Analysis of Children born with HIV at Bwaila Antenatal Clinic were analyzed. The chapter is presented in four sections, each contains the results associated with the research questions in the study.

The first section reports on the summaries of the data, identifying the basic approach in handling the Time series data has been discussed in the second section, the third section provides patterns and trend identification in the series. It also discusses ARIMA modeling approach and the fourth section reports on the ARIMA time series forecasting of children born with HIV.

The time series ARIMA models have been analyzed using a SAS program considering time in quarters as an independent variable and number of HIV infant births within the quarters as a dependent variable.

4.1 Summary Statistics

The figure 4.1 below provides a presentation of summary statistics for the examined HIV births of infants in the study.

The mean was found to be (6.2) within 32 observations and the maximum value of births recorded was 19. This means that there were approximately 6 births of children with HIV for every quarter since 2013 to 2020 on average.

Analysis Variable : HIV_births											
N	Mean	Median	Mode	Std Dev	Range	Skewness	Kurtosis	Lower 95% CL for Mean	Upper 95% CL for Mean	Minimum	Maximum
32	6.1875000	5.5000000	6.0000000	4.7684278	19.0000000	1.3897792	1.7043015	4.4682985	7.9067015	0	19.0000000

Figure 4.1: Summary Statistics on HIV births of infants

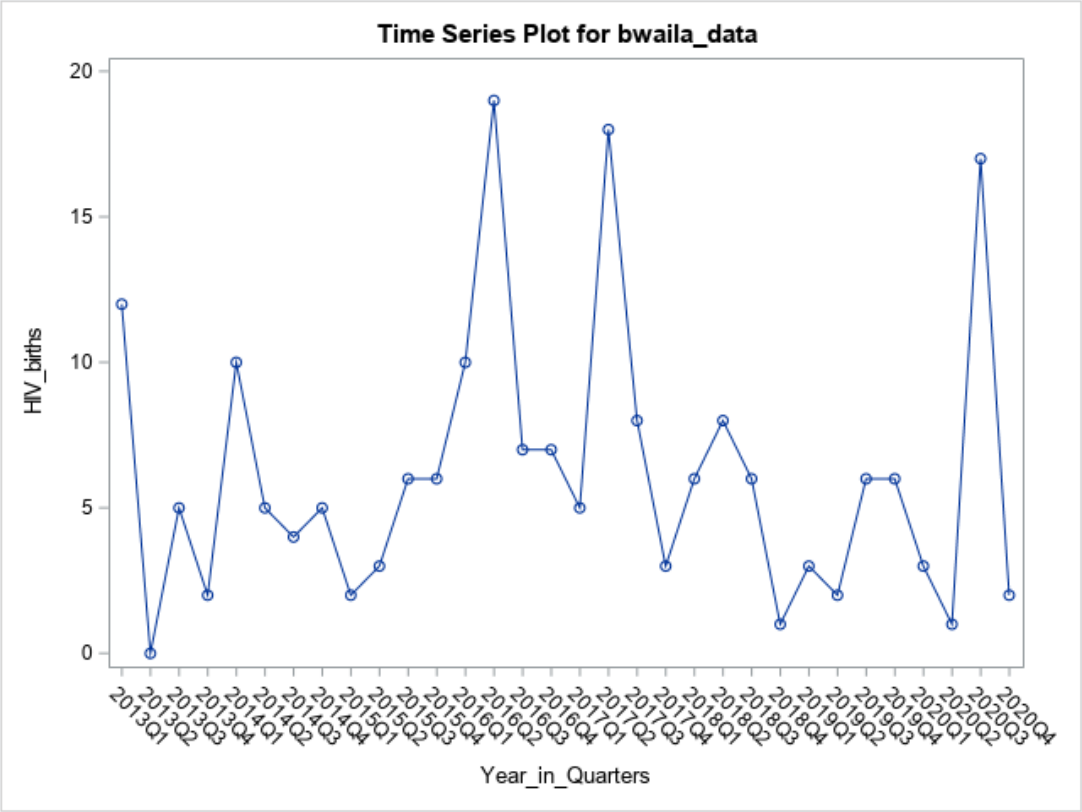


Figure 4.2: A Time Series Plot

4.2 Basic Approach in Time Series Identification of births

The subsequent figure 4.2 indicates the most basic approach technique applied in handling time series data, which involves plotting the Children born with HIV. The plot suggested an up and down effect with a considerable variation in the number of births in each quarter even though the numbers in births were very high in 2016 quarter 2 with (19) births, 2017 quarter 2 with (18) births and 2020 quarter 3 with (17) births respectively.

4.3 Pattern and Trend Analysis

4.3.1 ARIMA modeling Approach

The ARIMA Time series model involves a Box-Jenks approach for model identification, estimation and diagnostics. The Appendices B contains identifications and diagnostics of other models whose coefficients are indicated in table 4.1 that were used in selection of the best fit as well as an accurate model. From the Time series plot, the data were already stationary and therefore no differencing was performed in the analysis.

Augmented Dickey-Fuller Unit Root Tests							
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-12.2648	0.0101	-2.89	0.0052		
	1	-4.2163	0.1495	-1.33	0.1644		
	2	-2.3449	0.2853	-0.94	0.3009		
Single Mean	0	-31.5417	0.0002	-5.55	0.0002	15.46	0.0010
	1	-28.7271	0.0001	-3.55	0.0129	6.33	0.0149
	2	-29.9058	0.0001	-2.91	0.0565	4.24	0.0840
Trend	0	-31.5758	0.0003	-5.46	0.0006	14.92	0.0010
	1	-28.8786	0.0009	-3.49	0.0585	6.10	0.0780
	2	-30.2243	0.0004	-2.86	0.1901	4.09	0.3867

Figure 4.3: Test for Stationarity

ARIMA	(1, 0, 0)	(2, 0, 0)	(0, 0, 1)	(1, 0, 1)	(1, 1, 0)	(0, 1, 1)	(1, 1, 1)	(1, 1, 3)	(2, 1, 3)
Constants	6.19	6.19	6.19	6.19	-0.07	-0.11	-0.18	-0.19	-0.31
L1(AR)	-0.019	-0.019		-0.76	-0.52		0.18	-0.75	0.92
L2(AR)		0.007							0.07
L1(MA)			0.19	-0.71		0.79	0.93	-0.02	1.54
L2(MA)								0.6	-0.46
L3(MA)								0.3	-0.08
AIC	193.8	195.8	193.8	195.6	201.8	196.0	197.8	201.0	205.0
SBC	196.7	200.2	196.7	200.0	204.7	198.9	202.1	208.1	213.6

Table 4.1: ARIMA models

4.3.1.1 Identification

Pattern and Trend

Stationarity tests to identify the data with constant mean and variance were performed by running Augmented Dickey-Fuller Unit root test (ADF), where the computed Tau-Statistic (-5.55) with $Pr < Tau$ (0.0002) was statistically significant under the null hypothesis that the dataset was non-stationary. This is shown in figure 4.3. In the figure, the Trend test (-5.46) with $Pr < Tau$ (0.0006) was statistically significant indicating a Trend stationary of the series under H_0 : Non-trend stationary .

Partial Auto-Correlation Function and Auto-Correlation Function

Plots of the PACF and ACF were used to decide the Auto-regressive or Moving Average components. From the graph in figure 4.4 several models of ARIMA were assumed, the ACF was highly correlated at lag 1 and decays by the next lag where as the PACF cuts off in the first lag generating an ARIMA of the order (1, 0, 0). The frequency of the significant correlations in the ACF and the PACF indicated that the model generated had a non-seasonal pattern.

This model and other likely models generated through the process are furnished in Table 4.1.

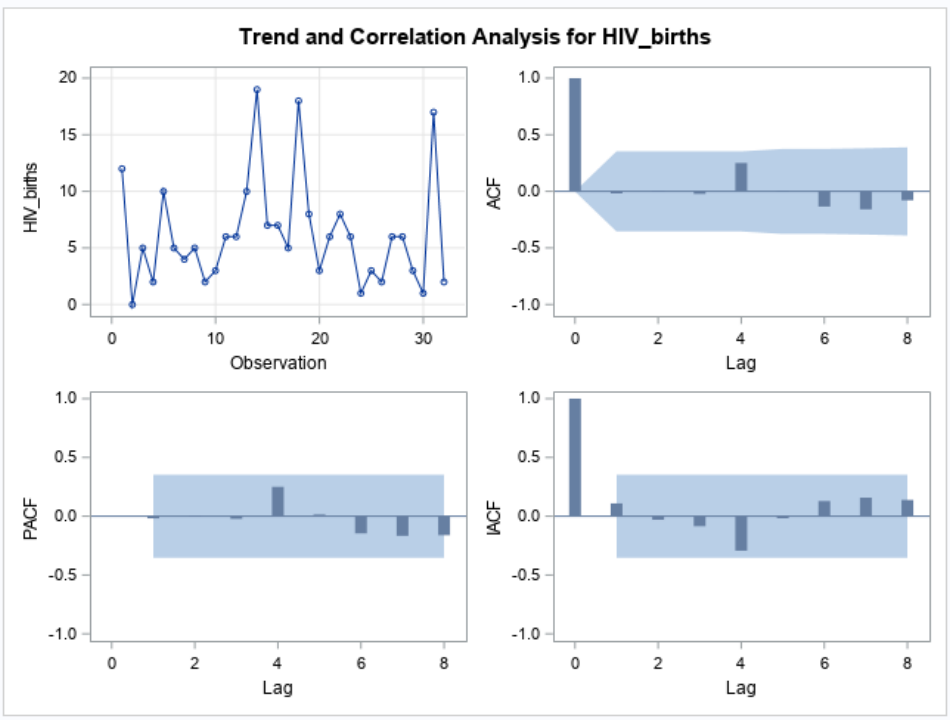


Figure 4.4: The ACF and PACF

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	3.23	6	0.7793	-0.018	0.004	-0.023	0.251	0.006	-0.135

Figure 4.5: White noise checks for the Residuals

4.3.1.2 Estimation Coefficient

The best fit coefficient of the ARIMA model was contained with its constant at ARIMA (1, 0, 0) in Table 4.1 at L1(AR), the coefficient was (-0.19). Theoretically presented as $\Theta_1(B) = (1+0.19B(1))$ recorded with a $Pr > 0.92$ which was also statistically significant. This had been obtained using Maximum likelihood Estimates.

4.3.1.3 Model Diagnostics Checking and Determining Accuracy of the Model

Checks for white noise

The model was checked for independency distribution of the residuals. This helps identification of misspecifications in the model under the null hypothesis that the data series is white noise or independently distributed, the chi-square distributions shown in the figure 4.5 recorded a $Pr < chisq(0.8)$ at 95% significance that concluded that a white noise process was present and the model was correctly specified since we did not reject the null hypothesis.

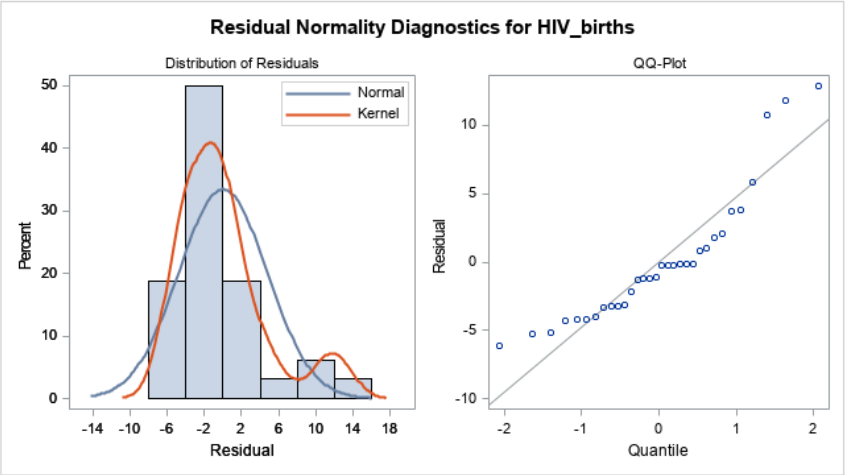


Figure 4.6: Normality tests

Goodness of fit

The Akaike Information Criterion (AIC) and the Schewatz Bayesian Information Criterion (SBC) were two measures of goodness of fit. They measure the trade-off between model fit and complexity of the model of the series data.

A lower AIC or BIC value indicates a better fit or a more parsimonious model, therefore with a lower AIC or SBC as indicated in Table 4.1 the parsimonious model was found to be ARIMA (1, 0, 0) in the study. The other way of doing this was checking the significance of the constant coefficient of the selected model, in this case from Table 4.1 looking at the coefficient at first Lag was not that closer to 1 thus AR1(1,0,0) (-0.019) hence significance was assured for parsimoniousness of the model.

Tests for Normality and Residuals

The properly fitted model of births was determined by a histogram plot of residuals distribution and the Quartile-Quartile plot for Normality of residual diagnosis. For the histogram in the figure 4.6, indicated normality distributions of the residuals with the highest mode and the spread of the residuals where the Q-Q plot on the other hand indicated a slight negative skewedness, therefore this provided an evidence that our model was well fitted by the normality assumption of the residuals.

4.4 Forecasting

The parsimonious model assured in the study, AR(1,0,0) had been considered the most accurate model through diagnostics observations in the approach. It had therefore been used to generate forecasts for 4 years, 2021 to 2025 quarters respectively with a 95% confidence interval that covered the actual predicted observations for the years in their preferred intervals as displayed in the figure 4.7.

Forecasts for variable HIV_births				
Obs	Forecast	Std Error	95% Confidence Limits	
33	6.2645	4.8464	-3.2343	15.7633
34	6.1852	4.8473	-3.3153	15.6857
35	6.1867	4.8473	-3.3138	15.6871
36	6.1866	4.8473	-3.3138	15.6871
37	6.1866	4.8473	-3.3138	15.6871
38	6.1866	4.8473	-3.3138	15.6871
39	6.1866	4.8473	-3.3138	15.6871
40	6.1866	4.8473	-3.3138	15.6871
41	6.1866	4.8473	-3.3138	15.6871
42	6.1866	4.8473	-3.3138	15.6871
43	6.1866	4.8473	-3.3138	15.6871
44	6.1866	4.8473	-3.3138	15.6871
45	6.1866	4.8473	-3.3138	15.6871
46	6.1866	4.8473	-3.3138	15.6871
47	6.1866	4.8473	-3.3138	15.6871
48	6.1866	4.8473	-3.3138	15.6871

Figure 4.7: 4 years forecast values for children born with HIV

The results from SAS procedure time series ARIMA forecast indicated that by the first quarter in 2021, the predicted births with HIV will be (6.2645); thus (6) in absolute terms and ranges within a 95% confidence limits (-3.2 to 15.8). The table further indicates the preceding forecasts for quarter 4 in 2025. These forecast values have been descriptively displayed by the chart in figure 4.8.

The forecasted values displayed in figure 4.7 indicated that there were flat forecasts, the confidence limits indicated the levels of minimum and maximum births expected, the fitted values indicated the exact point forecasts. The forecasts indicated no significant differences in births of infants through Mother to Child Transmission of HIV at the Hospital for the four years

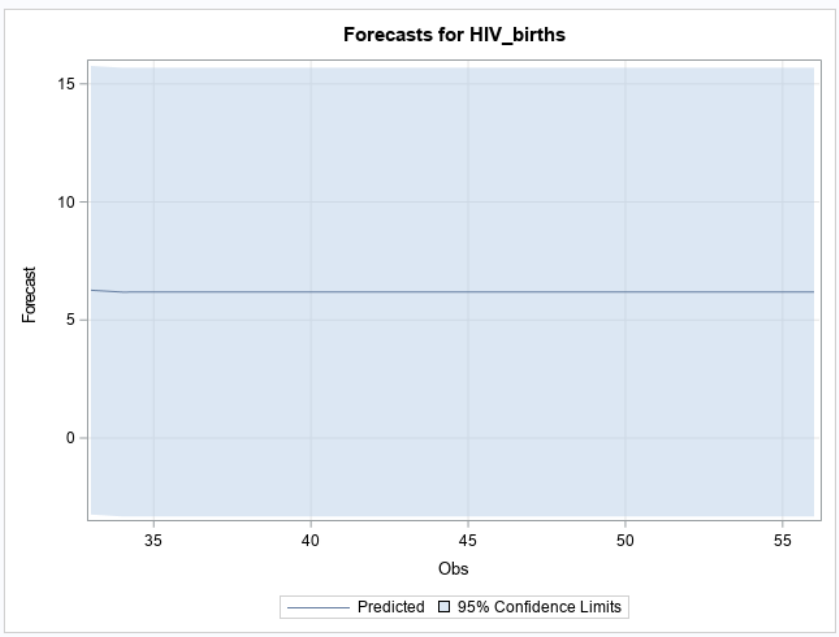


Figure 4.8: HIV births Forecasts

ahead. The four years of forecast were done with a consideration that, the lesser the forecast interval the more accurate the prediction will be, Douglas C.M et al (2008).

Chapter 5

Conclusion and Recommendations

The summary of the findings as well as the recommendations and conclusion suggested in this study were presented in this chapter.

5.1 Discussion

In the study and all the data used and analyzed, aimed to view children born with HIV considering a local phenomenon using time Series Analysis at Bwaila Antenatal Clinic in Lilongwe. With reference to the objectives of the study, the Box - Jenks ARIMA Time series performed accurately in all Box-Jenks approaches as outlined in chapter four of the study.

The study found out that, out of 32 observations included in the study, 6 Children were born With HIV in each quarter on average from 2013 to 2020. There was a considerable variation between births at each quarter, 2016 second quarter was recorded with highest number of births. The Time series plot indicated that despite the highest values of births in 2016, 2017 and in 2020, the number of births with HIV had been stabilized with a small variation (4.8) where the observations appeared within same limits horizontally over the period time 2013-2020.

The results further indicates that there was a non-seasonal pattern in the numbers of children born with HIV with a constant trend. The forecasts indicated that, for the four years to come, there will not be any decrease or an increase in births of infants with HIV. At least each quarter is associated with a birth of an HIV infant. This means that even though there exist a decrease in births of HIV infants at national level and other focus as depicted in the problem statement, this study contributes that there will not be termination of MTCT at a local perspective.

In line with the research questions which seek to identify the trend and patterns in HIV births of Children and identify an appropriate model for forecasting the Time series data at Bwaila. The results section had shown the patterns quarterly that, over the years from 2013 first quarter had a rising and dropping non-seasonal factor with a horizontal trend. The non-seasonal component in tells that the numbers of these births were not associated with the season of the

year in which they were born. And the trend meant a general direction in which these births are developing or changing, hence the series contains a stable and constant direction.

The appropriate and most accurate model for the forecast was identified by observing the significance of the coefficients and on the other hand through observing the lower value of the AIC and the SBC. Finally the forecasts indicated that HIV births were at drop slightly but there is never going to be a complete end locally. The reasons behind could be suggested as service delivery and adherence.

The results in chapter four strongly agrees with the findings obtained in previous research studies as depicted in the literature review that confirms a decrease in MTCT HIV births, although this study emphasized on a local application of Time series ARIMA on HIV births at Bwaila.

5.2 Recommendations

The study recommended the following suggestions towards the results.

Based on the conclusions in this study, practioners addressed in the significance should update the strategies and ensure penalties towards compliance, if it is to eliminate Mother to Child Transmission of HIV in the country and not only at Bwaila Antenatal Clinic in Lilongwe.

In line with the policies the government of Malawi articulated to terminate Mother to Child Transmission such as Promoting VCT for couples planning to have a child, and early attendance at an Antenatal Clinics and ensuring that HIV testing is routinely offered to all pregnant women attending Antenatal Clinics, the government through the Ministry of Healthy should propose an infant HIV diagnosis to be possibly performed as soon as the baby is born just as infants first vaccine is done to prevent categorizing errors arising from HIV birth through vertical transmission and through interactions together.

Though there is much secret towards knowing one's HIV status , it is worthy encouraging that residents on the other hand should push expectant comrades to report for check-ups at their nearest Antenatals, while NGOs play their part.

5.2.1 Recommendations to Further Research

To better understand the implications of the speculated results, future studies should be addressed on;

- Assessing Compliance towards HIV births through Mother to Child Transmission prevention measures, so as to observe how couples comply with the policies in the local

hospitals.

- Other studies could dwell on, Factors Affecting births of infants free from Infection, to determine what influences infants status locally.

5.3 Conclusion

Overall, it is clear from the findings of this study using the models of Time series on Children born with HIV at Bwaila Antenatal Clinic that HIV births are at a slight drop but there will never be an end toward it as depicted by the forecast in the next 4 years only if service delivery and adherence part is considered in another different way. The study used the ARIMA techniques to identify the pattern and provide a necessary forecast.

This study was therefore appropriate in the sense that it provided what is actually on the ground other than analyses performed globally, National wide Regional and for other collection of health clinics, with a Time series analysis consideration at Bwaila Antenatal Clinic.

Bibliography

[1] Bacon-Shone.J (2013). Introduction to Quantitative Research Methods, The University of Hong Kong, Hong Kong.

[2] Box, G. E. P and Jenkins, G. M (1976). Time Series Analysis Forecasting and Control. Oakland, CA: Holden Day

[3] Barbara G.T, Linda S.F (2007). Using Multivariate Statistics, 5^{TH} ED, pearson, Boston.

[4] Cohen, J., Cohen, P, West, S.G, and Aiken, L.S (2002). Applied multiple regression/correlation analysis for the behavioral sciences (3rd Edition). Mahwah, N.J: Lawrence Erlbaum.

[5] Davies, N, and Newbold, P (1979). "Some Power Studies of a Portmanteau Test of Time Series Model Specification," Biometrika 66(1), 153-155.

[6] Elliott, A.C, and Hynan, L.S (2007). "A SAS Macro Implementation of a Multiple Comparison Post Hoc Test." Proceedings of the Joint Statistical Meetings

[7] Government of Malawi National HIV/AIDS Policy :A CALL FOR RENEWED ACTION, Office of the President and Cabinet National AIDS Commission October 2003.

[8] Google_maps,jul-15-2021,7pm

[9] Malawi. Statistics [internet]. New York: United Nations Children's Fund; 2012.Available from: https://www.unicef.org/infobycountry/malawi-statistics.html116[cited2017Sep20]

[10] Malawi demographics profile 2017 [internet]. Charlotte: Index Mundi; 2017. Available from:https://www.indexmundi.com/malawi/demographics-profile.html[cited2017Sep20].

[11] Malawi Medical Journal : theJournalofMedicalAssociationofMalawi, 30Nov2017,(29(4):306-310DOI:10.4314/mmj.v29i4.5PMID:29963285PMCID:PMC6019542)

[12] Michele A.S et al, Evaluating the Impact of Prevention of Mother-to-Child Transmission of HIV in Malawi through Immunization Clinic-Based Surveillance, June 26, 2014 https://doi.org/10.1371/journal.pone.0100741

[13] M van Lettow et al Prevention of mother-to-child HIV transmission in Malawi, BullWorldHealthOrgan2018;96:256-265(doi:http://dx.doi.org/10.2471/BLT.17.203265)

[14] NCSS Statistical Software, NCSS.com

[15] SAS Institute, Inc. (2013). SAS/STAT software: reference, version 9.3. Cary, N.C.: SAS Institute, Inc.

[16] UNAIDS DATA 2020.

[17] UNAIDS 2016 Estimates, Malawi.

[18] UNAIDS unpublished HIV Estimates, 2012. Countdown To Zero: Elimination of New HIV Infections Among Children by 2015 And Keeping Their Mothers Alive. Unite For Children Unite Against AIDS, Malawi (draft).

[19] UNAIDS (1999) UNAIDS Report on the prevention of HIV transmission from mother to child strategic options. Geneva: UNAIDS Switzerland 1999.

[20] UNICEF: A case study on joint community-facility review of PMTCT dashboards in Malawi, February 2016

[21] World Health Organization Regional Office for Africa Brazzaville 2014, Implementation of Option B+ for Prevention of Mother-To-Child Transmission of HIV: Malawi experience

[22] Watts, D. (1997) Informed consent in emergency research: What every emergency nurse should know, JournalofEmergencyNursing,23(1):70-4.

[23] Yaffee.R, McGee.M. Introduction to Time series and forecasting with applications of SAS and SPSS, Academic Press INC, New York.

Appendix A

Appendix Data Collection Letter

Ref. No.:
Telephone No.: 265 726 466/464
Telefax No.: 265 727817
Telex No.:
E-Mail. lilongwedho@malawi.

COMMUNICATIONS TO BE ADDRESSED TO:

In reply please quote NO DZH/MALAWI,
Lilongwe District Health Office
P.O. Box 1274
Lilongwe
Malawi

5th May, 2021

The In-Charge : HMIS office

Dear Sir/Madam

PERMISSION TO CONDUCT RESEARCH IN LILONGWE DISTRICT

Permission has been granted to the bearer of this letter **Joachim Kapalamula–** Pursing a Bachelor of Science in Applied Statistics from Catholic University to conduct research in Lilongwe District.

RESEARCH TITLE:-

"TIME SERIES ANALYSIS OF CHILDRED BORN WITH HIV THROUGH MOTHER CHILD TRANSMISSION AT BWAILA ANTENATAL CLINIC"

Any assistance rendered would be appreciated.

MINISTRY OF HEALTH
LILONGWE DISTRICT HEALTH OFFICE
DMO/MO
05 MAY 2021
P.O. BOX 1274
LILONGWE

Duncan Banda
For : DIRECTOR OF HEALTH AND SOCIAL SERVICES -LL

Appendix B

Appendix SAS Outputs

B.1 SAS outputs for other Models

Model diagonistics, Normality and Residual checks for various models as used in the study

B.1.1 ARIMA model (2,0,0)

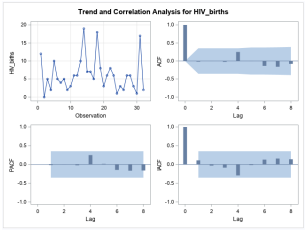


Figure B.1: Trend and Correlation ACF and PACF

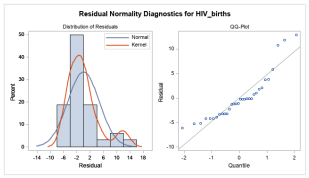


Figure B.2: Residual Normality diagonistics

Model diagonistics, Normality and Residual checks for ARIMA (2,0,0)

B.1.2 ARIMA Model(1,0,1)

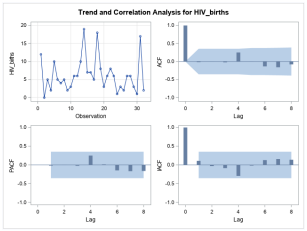


Figure B.3: Trend and Correlation ACF and PACF

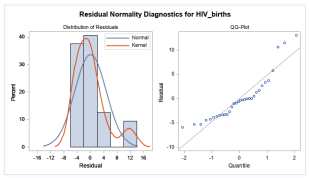


Figure B.4: Residual Normality diagonistics

Model diagonistics, Normality and Residual checks for ARIMA model (1,0,1)

B.1.3 ARIMA model (1,1,0,)

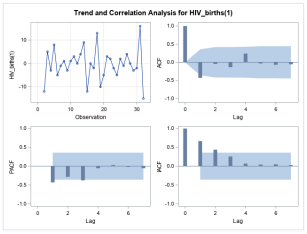


Figure B.5: Trend and Correlation ACF and PACF

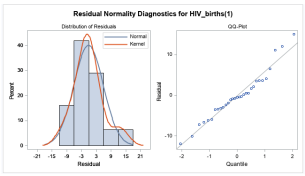


Figure B.6: Residual Normality diagonistics

Model diagonistics, Normality and Residual checks for ARIMA model (1,1,0)

B.1.4 ARIMA model (0,1,1,)

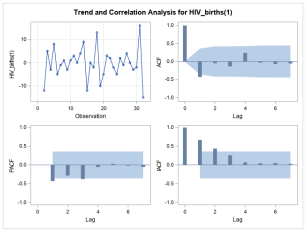


Figure B.7: Trend and Correlation ACF and PACF

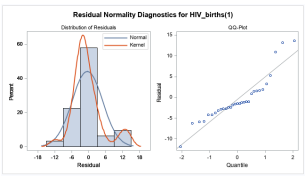


Figure B.8: Residual Normality diagonistics

Model diagonistics, Normality and Residual checks for ARIMA model (0,1,1)

B.1.5 ARIMA model (1,1,1,)

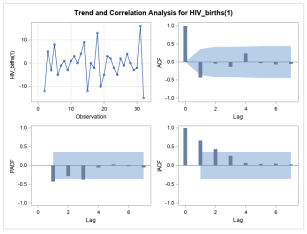


Figure B.9: Trend and Correlation ACF and PACF

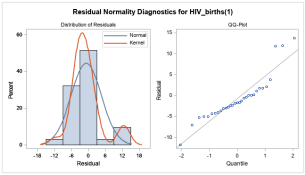


Figure B.10: Residual Normality diagonistics

Model diagonistics, Normality and Residual checks for ARIMA model (1,1,1)

B.1.6 ARIMA model (1,1,3,)

B.1.7 ARIMA model (2,1,3,)

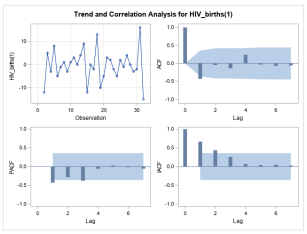


Figure B.11: Trend and Correlation ACF and PACF

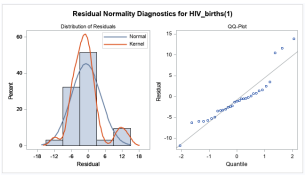


Figure B.12: Residual Normality diagonistics

Model diagonistics, Normality and Residual checks for ARIMA model (1,1,3)

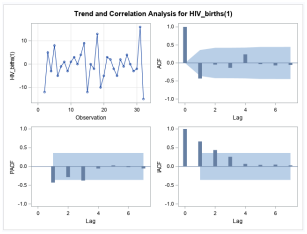


Figure B.13: Trend and Correlation ACF and PACF

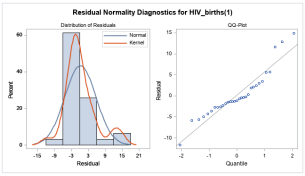


Figure B.14: Residual Normality diagonistics

Model diagonistics, Normality and Residual checks for ARIMA model (2,1,3)

Appendix C

Appendix SAS Commands

/* Time Series Analysis of Children born with HIV Command Exercutions */
/*The Plotting of the Time series data */
/*Joachim C.kapalamula */

Title 'Bwaila Time series dataset';
Data bwaila_data;
input_Date $ *HIV_births*;
datalines;
2013Q1 12
2013Q2 0
2013Q3 5
2013Q4 2
2014Q1 10
2014Q2 5
2014Q3 4
2014Q4 5
2015Q1 2
2015Q2 3
2015Q3 6
2015Q4 6
2016Q1 10
2016Q2 19
2016Q3 7
2016Q4 7
2017Q1 5
2017Q2 18
2017Q3 8
2017Q4 3
2018Q1 6
2018Q2 8

```
2018Q3 6
2018Q4 1
2019Q1 3
2019Q2 2
2019Q3 6
2019Q4 6
2020Q1 3
2020Q2 1
2020Q3 17
2020Q4 2
;
run;

    Title 'Data print';
proc print data = bwaila_data;
run;

    * Creating a differenced variable;

    data bwaila_data; /*some of the variables used*/
set bwaila_data;
HIV = dif(HIV_births);
lHIV = lag(HIV_births);
ldHIV = lag(HIV_births);
run;

    %let ylist = HIV_births;
%let dylist = HIV;
%let time = t;
%let lylist = lHIV;
%let trend = trend;

    Title 'Summary statistics'; /*Summary statistics*/
proc means data = bwaila_data n mean median mode std range skewness kurtosis clm min max;
var HIV_births;
run;

    Title 'Time Series Plot for bwaila_data'; /*Time series Data plot*/
proc sgplot data = bwaila_data;
series x = Date y = HIV_births/markers;
```

```
run;

    Title 'ARIMA identification'; /*ARIMA approach 'Identification stage' and stationarity test
*/
proc arima data = bwaila_data;
identify var = HIV_births stationarity = (adf);
run;

    Title 'ARIMA(1,0,0) or AR(1)'; /*Various model estimations*/
proc arima data = bwaila_data;
identify var = HIV_births;
estimate p=1 method=ml;
run;

    Title 'ARIMA(2,0,0) or AR(2)';
proc arimadata = bwaila_data;
identify var = HIV_births;
estimate p=2;
run;

    Title 'ARIMA(0,0,1) or MA(1)';
proc arima data = bwaila_data;
identify var = HIV_births;
estimate q=1;
run;

    Title 'ARIMA(1,0,1) or ARMA(1,1)';
proc arima data = bwaila_data;
identify var = HIV_births;
estimate p=1 q=1;
run;

    Title 'ARIMA(1,1,0)';
proc arima data = bwaila_data;
identify var = HIV_births(1);
estimate p=1;
run;

    Title 'ARIMA(0,1,1)';
proc arima data = bwaila_data;
```

```
identify var = HIV_births (1);
estimate q=1;
run;

    Title 'ARIMA(1,1,1)';
proc arima data = bwaila_data;
identify var = HIV_births (1);
estimate p=1 q=1;
run;

    Title 'ARIMA(1,1,3)';
proc arima data = bwaila_data;
identify var = HIV_births (1);
estimate p=1 q=3;
run;

    Title 'ARIMA(2,1,3)';
proc arima data = bwaila_data;
identify var = HIV_births (1);
estimate p=2 q=3;
run;

    Title 'ARIMA (1,0,0) forecasting'; /* ==========forecasting==========*/
proc arima data = bwaila_data;
identify var = HIV_births;
estimate p=1;
forecast lead=16;
run;
```