2019

A Bayesian approach to Geo spatial analysis of HIV viral load data

Joy H. M. Kareko
Strathmore Institute of Mathematical Sciences (SIMS)
Strathmore University

Recommended Citation

Kareko, J. H. M. (2019). *A Bayesian approach to Geo spatial analysis of HIV viral load data* (Thesis, Strathmore University). Retrieved from http://su-plus.strathmore.edu/handle/11071/6737

A Bayesian Approach to Geo Spatial Analysis of HIV Viral Load Data

Kareko Joy Hilda Mukami

Master of Science in Statistical Science

2019

A Bayesian Approach to Geo Spatial Analysis of HIV Viral Load Data

Kareko Joy Hilda Mukami

Submitted in partial fulfillment of the requirements for the Degree of Masters of Science in Statistical Science at Strathmore University

Strathmore Institute of Mathematical Sciences
Strathmore University
Nairobi, Kenya

June 2019

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the dissertation contains no material previously published or written by another person except where due reference is made in the dissertation itself.

Kareko Joy Hilda Mukami

6th June 2019

Approval

The dissertation of Kareko Joy Hilda Mukami, was reviewed and approved by the following:

Dr.Collins Odhiambo

Lecturer, Strathmore Institute of Mathematical Studies

Strathmore University

Ferdinand Othieno

Dean, Strathmore Institute of Mathematical Studies

Strathmore University

Professor Ruth Kiraka

Dean, School of Graduate Studies

Strathmore University

Abstract

HIV is currently ranked among the leading causes of death in Kenya and in the world, with an estimated 1.5 Million Kenyans living with HIV and 28,000 deaths recorded annually as a result of AIDS related illnesses. In 2014, UNAIDS launched a 90-90-90 strategy the aim was to diagnose 90 per cent of all HIV- positive persons, provide antiretroviral therapy (ART) for 90 percent of those diagnosed, and achieve viral suppression for 90 per cent of those treated by 2020.

This study is motivated by the need to assess the 3rd 90; viral suppression for 90 per cent of those ART treated and seeks to analyze one statistical paradigm (Bayesian) that have conventionally been used for geospatial trends. Use of Bayesian approach has been used previously to assess the prevalence and incidence of diseases however, this dissertation seeks to evaluate Bayesian Approach to spatial trends of HIV Viral Load Suppression in Kenya. We revisit the theoretical framework of the Bayesian Approach and apply real data from the Kenyan setting spanning from 2012 to 2017.

Results show Bayesian Approach to be robust, in depth and entails more information when modelling spatio-trends of Viral Load suppression. Further, First Line ART regimen, HIV-TB co-infection and retention rates are significant predictors of Viral Load suppression spread.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AIDS	Acquired Immune Deficiency Syndrome
AIC	Alkaike Information Criterion
ART	Antiretroviral Therapy
ART-T	Antiretroviral Therapy Treatment
CAR	Conditional Autogressive Models
DIC	Deviance Information Criterion
GIS	Geographical Information Systems
GMSM	Gay and other Men who have Sex with Men
HIV	Human Immunodeficiency Virus
LMICs	Low and Middle-Income Countries
NACC	National Aids Control Council
NASCOP	National Aids and STI Control Program
PLWHIV	People Living With HIV
PWID	People Who Inject Drugs
PMTCT	Prevention of Mother to Child Transmission
SMR	Standard Morbidity Ratio
UNAIDS	United Nations Program on HIV/AIDS
VL	Viral Load
WHO	World Health Organization

Acknowledgements

This thesis would not have been successful without strength, good health, wisdom and a good support system which i owe to God everyday. Its surely a dream come true.

The constant encouragement,criticism and help from my Supervisor, Dr. Collins Odhiambo, helped me to discover hidden strength and capacity in me that i never thought i had. This helped me to find courage, think broadly and widely and still meet strict deadlines that would not have been easy to meet except for the daily reminders and push. Any student who juggles a full time job and school would clearly understand that its not easy.

The team at Strathmore Institute of Mathematical Sciences was great as well, as they provided great learning material, support, well knowledgeable lecturers and great seminars that assisted us have a better perspective of our career journeys.

Lastly, i am grateful to my family and my friends as well as my classmates, for the constant support, and checking, ensuring that i prioritized school and continued to work hard.

Dedication

This dissertation is dedicated to my mother and my brother who made me realize that we can achieve anything we task our minds to do, through prayer, consistency, persistence, focus, proper planning, some fun and a healthy dose of humour and my grandfather, who constantly pushed me to go back to school and pursue a masters degree because "learning never ends" and "good brains shouldn’t go to waste".

Chapter One

Introduction

This chapter is a discussion on the HIV trends and statistics from a global perspective as well as a Kenyan setting.

It also shows the gaps that this study seeks to address; as well as the research objectives; the scope and the significance of this study.

1.1 Background Information

In 2014, United Nations Program on HIV/AIDS (UNAIDS) launched the 90-90-90 target goals whose aim was to achieve 90 per cent suppression among the HIV infected persons , to ensure that 90 per cent of the HIV infected people are under ARV treatment, and ensure that 90 per cent of all HIV infected persons have been diagnosed and are therefore aware of their status by 2020 as per (Levi et al., 2016).

Globally; an estimated 36.7 million people are living with HIV, of which 2.1 million are children. This translates to 1.8 million new infections each year (approximately 5000 new infections daily), of which about 160,000 are children below the age of 15 years as shown by (Sidibé et al., 2016). In terms of prevalence; Swaziland is currently leading with 27.2 per cent, but the numbers are especially high in South Africa, where about 7.1 Million people are living with HIV and 3.2 million in Nigeria, that is 18.9 per cent and 2.9 per cent prevalence respectively (Cheluget et al., 2006).

It is estimated that about 1.5 million Kenyans were living with HIV in 2016 as per a report done by (Sidibé et al., 2016) whereas an estimated 71,034 new infections were recorded among adults (Age 15+), and 6,613 new infections recorded among children (Ages 0-14) in that same year.

This translates to a prevalence rate of about 7 per cent for women, who are most vulnerable to HIV infections and 4.7 per cent for men, with the most affected county being Homabay County; whose prevalence is 26 per cent. HIV therefore contributes to 29 per cent of annual adult deaths, 20 per cent of maternal mortality and 15 per cent of deaths of children under the age of five years. This simply means, for every 100 adult deaths, 29 of them are due to AIDS related illnesses.

Treatments for HIV involves a combination of different drugs, which is the ARVs-T, which prevents the virus from replicating; maintaining the immunity levels of the person while slowly reducing the progression of the virus in the body. According to a study done on GMSM in Australia in 2012, 50-70 per cent of people diagnosed with HIV and were receiving treatment through administration of ARV-T; 85-95 per cent had a suppressed viral load (Nicole et al., 2015). Viral suppression rates have also been found to be low among children on ART in LMICs; even after continuous ART administration and periodic monitoring of the same a reported by (Boerma R and J, 2016).

According to another report conducted in the USA, 85 per cent of the people diagnosed with HIV; 49 per cent of them had the virus under control through ART (CDC, 2005).

Administration of ARV reduces the viral load content per millimeter of blood suppressing the virus, hence slowing down progression of HIV to AIDS (Ickovics and Meade, 2002) and this can prevent transmission of the HIV virus to uninfected partners (Cohen et al., 2011). Since HIV virus attacks the immune system cell in the body, known as CD4 helper lymphocyte cells, making the body unable to fight other infections; administration of ART will almost successfully suppress the virus, reducing it to a certain level even undetectable. Once the CD4 cell count falls below 200 cells/mm3, HIV is said to have progressed to AIDs.

Studying retention in HIV care improves survival and HIV Viral control and decreases race or ethnicity related healthcare disparities (Tobias et al., 2007) and is an important indicator for quality of healthcare services (Braitstein et al., 2010). Its is however estimated that only 75 per cent of HIV-diagnosed patients are linked to care and only 66 per cent of them are retained in care successfully in the United States(Gardner et al., 2011). In Kenya, 64 per cent of adults were on ART as compared to 65 per cent of children in 2016 (UNAIDS, 2017)

In 2015,81 per cent of people initiated on treatment were still in care after 12 months, which is an improvement on the 2013 retention rate which was 70 per cent (NACC, 2016)

This study will therefore aim at finding out the suppression rates among children and adults and also establish the percentages of this population who have the virus under control due to the administration of the ART. This study will also show the trends in the suppression rates while trying to model the same for future use.

1.2 Problem Statement

HIV is a continued concern for many governments. Reduction of new annual infections will reduce the HIV prevalence rates in the country.

Periodic monitoring of the viral loads among the infected assists in proper administration of ART as well as prevent transmission of HIV from mother to infant during birth or maternal care and even between partners.

To establish how well the ART is working among demographic population; there is need for us to examine the suppression rates and patterns of the HIV virus among demographic population, the children and the adults as well as model these trends; to have a better understanding of how the ART is working in that demographic group.

High suppression rates may indicate that the patients are keen to follow the prescribed dosages for the ART while the opposite will also be true.

Whereas GIS has been used to model epidemic spread of diseases no attempt has been made to extend its application in critical disease programming and monitoring, specially in viral load monitoring in HIV/AIDS.

1.3 Research Objective

The aim of this study is to model distributions and patterns of viral load suppression using Geographic Information Systems as well as use of Bayesian Statistics to model outcomes of Viral Load suppression.

1.3.1 Specific Objective 1

To conduct location analysis by focusing on Viral Load analysis

1.3.2 Specific Objective 2

To conduct analysis of Viral Load patterns and HIV predictors with intention of analyzing health complexities.

1.4 Scope of the Study

This study analysed trends and patterns in HIV viral suppression rates and HIV retention among demographic population in the 47 counties of Kenya through use of GIS from Jan 2012 to December 2017.

1.5 Significance of the Study

This study will assist the government in decision making on distribution of resources to areas with least suppression rates.

This study will also be beneficial to the National Aids and Control Council (NACC)who are currently mandated with the task of reducing HIV prevalence rates in the country through education, training and awareness, and rigorous administration of ARVs-T and to achieve World Health Organization ambitious target of achieving 90 per cent viral suppression for HIV patients on ARV-T.

Globally; this research will assist in continuous improvement of Antiretroviral Therapy; given the current concerns on mutation of the HIV virus which might render the current HIV drugs irrelevant.

Chapter Two

Literature Review

This chapter provides discussions on HIV trends, prevalence, and viral suppression rates both globally and locally, as well as new developments in HIV.

2.1 HIV Trends and Statistics

2.1.1 A global look on the trends and statistics

Overall, an estimated 36.9 million people were living with HIV at the end of 2017. Among these people, 1.8 million are children. Majority of these people live in sub-Saharan Africa. 19.4 million are people living in East and Southern Africa, 6.1 Million are living in western and central Africa, 5.1 Million live in Asia, 2.1 are living in West and Central Europe and North America, another 2.1 live in Latin America and the Caribbean, 1.6 Million are in East Europe and Central Asia, while only 230,000 people are in the middle east and North Africa (Sidibé et al., 2016).

In terms of mortality, 940,000.00 died of HIV related illnesses worldwide, while 1.8 million people became newly infected, in 2017. Since the start of the epidemic, 77.3 million have become infected; while 35.4 million have died due to AIDs related illnesses (Sidibé et al., 2016). This is a reduction in AIDs related illnesses by 51 per cent since 2004, while the number of new infections declined by about 16 per cent among adults and 35 per cent among children since 2010.

Over time, disease monitoring indicators of HIV have changed from CD4 count to Viral load. Viral load measures the amount of virus per milimeter of blood. In order to measure progress, there is need to monitor VL in PLWHIV and assess the effectiveness of the ART-T to them.

2.1.2 Trends and statistics in Kenya

In Kenya,an estimated 1.6 million people are living with HIV. According to a report released in 2017, 36,000 people died from AIDS-related illnesses in 2016. Kenya's HIV epidemic is generalized, and therefore affects all groups of people,from children, young adults, men and women alike. (Sidibé et al., 2016)

Groups most affected by HIV are sex workers, who have the highest prevalence (NACC, 2016) at 29.3 per cent prevalence. Prevalence among men who have sex with other men is at 18.2 per cent, while People Who Inject Drugs are estimated to be 18.3 per cent were living with HIV. Further; its estimated, that young women are more likely to acquire HIV than their male counterparts,and accounted for 33 per cent of all new infections, as compared to 16 per cent of new infections in men (NACC, 2016). The young adults are more vulnerable to new infections due to factors such as incorrect perception of the risk on exposure, forced sex and sexual violence among others.

In Kenya, measurement of VL is a great way of monitoring progress in the HIV epidemic, as well as assessment of retention and therefore adherence. Over the years, the prevalence rates may have declined but there is need to assess the progress given the 90-90-90 targets by UNAIDS and WHO (Levi et al., 2016).

2.2 Access to treatment - Global and group perspectives

Continuous education and HIV awareness has greatly increased knowledge of HIV in schools, as it forms part of the curriculum in all schools. Mother to Child Transmission has also been greatly reduced. Number of newly infected children fell from 12,000 to 6,600 in 2015, attributed majorly to PMTCT services (UNAIDS, 2017).

It is estimated that 64 per cent of adults and 65 per cent of children were accessing ART in 2016 as a result of increased HIV awareness.

In 2015, 81 per cent of people on treatment, of whom 64 per cent were virally suppressed, which equates to 51 per cent of all people in Kenya living with HIV (UNAIDS, 2017).

HIV is a leading cause of death among young adults,mostly because most young people will find it difficult to adhere to treatment, leading to a form of drug resistance, making HIV treatment options more limited with time (Onywera et al., 2017).

A 2015 report by the World Health Organization estimates global coverage on ART-T to be 46 per cent, this is estimated to be 17 million people. The report attributed the major increase to coverage in Africa regions where coverage increased from 24 per cent in 2010 to 54 per cent in 2015.

Global coverage of ART-T is a good indicator of having HIV under control, however, the expected target should be 90 per cent and therefore more need to monitor progress of the disease.

2.3 Suppression rates among demographic population

Increasing viral suppression to a low or undetectable level improves the health of of those living with HIV and prevent sexual transmission of HIV to partners (Doshi et al., 2017).

In 2014, 57.9 per cent of people living in the US were virally suppressed from 54.7 per cent in 2013. Among the youth, 48.1 percent were virally suppressed, while 50.3 percent of PWID were also virally suppressed (Doshi et al., 2017).

In Kenya, the national viral load monitoring program is used to monitor viral load across the country, with over 2,000 facilities currently, and over 40,000 tests conducted each month. According to NASCOP,there are 871,156 suppressed patients while 137,452 are not suppressed, this is 86.4 per cent as compared to 13.6 per cent respectively.

The national viral load monitoring program monitors VL suppression,however, there may be other indicators of VL suppression such as retention in care and therefore adherence,availability of donor funding to specific areas,HIV co-infections such as TB, Treatment level Regimens(whether 1st Line or 2nd or 3rd Line) and many more.

2.4 Use of Geographical Information Systems in Disease Mapping

GIS has been used extensively in epidemiology for disease surveillance and intervention monitoring. Through mapping, governments have been able to identify disease spread and monitoring of the same, leading to proper control of epidemics such as trypanosomiasis in Africa (Clarke et al., 1996).

GIS was particularly used for planning HIV prevention interventions for High Risk Youth, whereby the information was used to map neigbourhoods where the youths were at highest risk of HIV/AIDS contraction and thereafter helped in decision making (Geanuracos et al., 2007).

GIS was also used to map HIV in India, to help have a perspective of the disease spread and monitoring (Kandwal et al., 2009)

GIS was also used to show the spatial spread of the AIDS epidemic in Ohio (Casetti and Fan, 1991).

GIS however, has not been extensively used in the area of disease programming and disease monitoring indicators for HIV which is the objective of this study.

2.5 New developments in HIV Treatment

There has been new developments in HIV recently especially in treatment of HIV and prevention such as HIV vaccines whose research is currently ongoing, with good progress being made, however, great challenges still remain on the same, such as a new animal model data that indicates that the viral reservoir is seeded before plasma virus becomes detectable in acute infection, (Sidibé et al., 2016) hence showing the vaccine might not be as effective as anticipated.

There are other developments as well on developing the cure for HIV with several methods being modelled and developed as outlined by Clara Fernandez of Labiotech.

These methods include, Stopping the replication of HIV, a method used to also treat herpes, Shock and Kill, which generally targets the hidden HIV reservoir that contains the inactive viruses that are waiting to be active and replicate, Immunotherapy, where the immune cells are supercharged and strengthened to be able to attack the virus itself and lastly gene therapy, which is a method that currently seeks to edit the DNA of a person and introduce a mutation to stop HIV. This is because, an estimated 1 per cent of the world is naturally immune to HIV.

2.6 Conclusion

Geographic Information Systems have not been extensively used in the area of disease programming for HIV and has only been used to map diseases and show spread of outbreaks but has not been used to show certain aspects of diseases such as retention, viral load monitoring and such other aspects useful in disease monitoring hence the gap that has been looked at in the study.Further, no attempt has been made to use bayesian approaches on such disease monitoring indicators such as VL suppression.

Chapter Three

Methodology

This chapter contains the methods that were used to achieve the outcomes of interest in this study.

3.1 Study Design

Spatial analysis was used in this study. The study population involved all HIV positive patients who were actively on ART from 2014 to 2017 as per NASCOP VL website *(http://viralload.nascop.org/)*. Participating sites included all facilities registered by government of Kenya with MFL(Master Facility List) codes. Abstracted data was assembled into a uniform excel format disaggregated by counties. The viral load data consisted of 3123 sites that treated HIV adult and pediatric patients. The data collected also included other variables i.e. testing month, redrawn samples, 1st line patients, patients with HIV/TB co-infection and retention rate from January, 2014 to December, 2017 were extracted from the data warehouse.

Areas of interest were as listed below;

1. Current suppression rates in Kenya, in the different counties in Kenya.
2. Current Retention rates in Kenya, at county level.
3. Is there any correlation between retention and suppression rates.
4. Is there any relationship between Patients on first line treatment, HIV-TB co-infection rates and VL suppression rates.
5. Reason for low/high suppression rates and retention rates
6. Comparison of the current suppression rates as compared to the expected rates of suppression.

3.2 Study Area and Data Collection

The study area was the 47 different counties in Kenya. Data was obtained from health centers at a county level which comprised of data including viral load tests, viral load results with $>$ 1,000 mL/copies were collected retrospectively, by electronic abstraction from each site. Abstracted data were sent in electronic excel format to a Masterfile.

Electronic data received was reviewed to ensure that each data element was correctly formatted and that all elements were captured. Data elements with incorrect formatting, unknown or incomplete information, or other inaccuracies were reviewed with the site and corrected. The data was combined across sites to achieve a uniformly constructed multi-site database.

3.3 Statistical Analysis

Data was examined for viral load tested for patients on ART and virologic results were used to determine VL suppression. The primary area of interest was to analyze patterns of viral load suppression over the years. Given that the study sample is drawn from all facilities providing clinical care, individuals who were being initiated ART are likely to have VL test together with those already on ART.

The data also included information on retention rates, HIV-TB Co-infection rates, as well as Number of patients on 1st line treatment and was used to examine whether there was any relationship between the same. Comparisons of hazards in non-referent categories were performed with linear combinations. For each interaction,effects of enrollment year on the hazard within categories of the other variables were assessed. To interpret interactions, we estimated the linear year trend (slope) of the hazard within categories of the other variable. Cumulative incidence curves were generated to display the cumulative incidence of ART initiation and virologic suppression prior to the competing risks by enrollment year.

A non-parametric cumulative incidence function was used to estimate cumulative incidence in a competing risk situation. Cuzick's test was used for non-parametric trends . Statistical analysis was performed using R Studio version 3.5.3.

3.4 Data Exploration and Visualization

Data Exploration and Visualization involved getting a summary of the variables that were of interest in order to scrutinize the patterns that emerged.
The data contained 10 variables. The variables were basically,County, No. of Facilities Sending Samples, Rejected Samples, Redrawn Tests, Tests, Results with VL $>$ 1000,Patients on 1st Line treatment,Retention Rates, Patients with HIV-TB Co-infection and Year in which these samples were obtained.

An additional variable which was of interest was obtained by subtracting the Results with VL $>$ 1000 from the total Tests obtained and then finding a percentage of the same so as to obtain the variable of individuals who were virally suppressed but as a percentage. This was done for each county and then done for each year from 2012 to 2017.

Choropleth maps which show information by coloring each component area with color, providing an indication of the magnitude of the variable of interest which were then used as a means for visualizing this data over the same period of time.

3.5 Relative Risk Estimation

The data was from the 47 counties. Total Number of Tests taken were used interchangeably to denote the total population of that region.

Let N_i, denote the total number of samples taken in region i and let Y_i, denote the observed number of suppressed cases in region i. The total number of tests taken in region i is given by

$$N_i = \Sigma_{i=1}^{J} N_i. \tag{3.1}$$

The total number of cases of people who are suppressed in region i is given by

$$Y_i = \Sigma_{i=1}^{J} Y_i. \tag{3.2}$$

The total number of cases in the population is given by

$$N_+ = \Sigma_{i=1}^{J} \Sigma_{i=1}^{J} N_{ij}. \tag{3.3}$$

Hence the expected number of cases in region i is given by

$$E_i = r_+ N_i. \tag{3.4}$$

where $r_+ = \frac{Y_+}{N_+}$.

The Standard Morbidity Ratio(SMR)(Waller and Gotway, 2004) is a ratio obtained to show an estimate of the risk and is obtained by comparing the observed number of cases with the expected number of cases while mapping a disease in an area. This ratio will be obtained in this study and used in this study to show the number of expected virally suppressed population and the observed cases. The ratio is given by the below;

$$SMR = \frac{Y_i}{E_i}. \tag{3.5}$$

Values greater than one imply a suppression rate of HIV is higher than expected, while below one indicate a rate lower than expected for the respective area. However, a low value of E_i can happen if the population of a certain place is too low or if the disease under study is rare.To overcome this problem, Bayesian hierarchical spatial models are adopted. The study embraced use of the Bayesian Approach.

3.6 Moran's Index Statistic

Moran's I (Moran, 1950) is used to test the hypodissertation that there is no spatial autocorrelation in the outcome variable.It measures the overall spatial autocorrelation that exist in the data under study. High Moran's Index will indicate that the data is highly correlated while the opposite is also true. The expected value of Moran's I is $-1/(N-1)$. Values of Moran's I that exceed $-1/(N-1)$ indicate positive spatial autocorrelation, in which similar values, either high values or low values are spatially clustered.

Values of I below $-1/(N-1)$ indicate negative spatial autocorrelation, in which neighboring values are not similar.(Li et al., 2007)

3.7 Bayesian Approach

The study used Conditional Autoregressive Methods in this approach.CAR models propose conditionally autoregressive priors (CAR priors) in an empirical Bayes setting instead of the joint prior distribution.In addition, Breslow and Clayton (Breslow and Clayton, 1993)apply CAR priors as random effects distributions within likelihood approximations for Generalized Linear Mixed Models. CAR models have a specified conditional mean and variance as shown below;

$$E[Y(S_i)|Y_{-i}] = X(S_i)'\beta + \Sigma_{j=1}^{N} cij[Y(S_j) - X(S_i)'\beta]. \tag{3.6}$$

$$Var[Y(S_i)|Y_{-i}] = \delta_i^2, i = 1, ..., N. \tag{3.7}$$

where the c_{ij} denote spatial dependence parameters.The primary purpose of CAR models is to provide a modelling mechanism to account for residual spatial correlation not explained by spatial patterns in covariate values. A number of different conditional autoregressive prior models have been proposed in a disease mapping context and in this case disease monitoring indicators context as below;

1. **Intrinsic Model.**
   This is one of the simplest CAR model which was proposed by (Besag and Higdon, 1999) and (Besag et al., 1991)

   It has full conditional distributions given by;

   $$\phi_k|\phi_{-k}, W, \tau_1^2 \sim N\left(\frac{1}{n_k}\sum_{j\sim k}, \phi_j, \frac{\tau_1^2}{n_k}\right). \tag{3.8}$$

   This equation is interpreted as follows; The conditional expectation of ϕ_k is equal to the mean of the random effects in neighbouring areas, while the conditional variance is inversely proportional to the number of neighbours n_k.
   This variance structure recognises the fact that in the presence of strong spatial correlation, the more neighbours an area has the more information there is in the data about the value of its random effect.

   The advantage of this model is that it is simple and easy to apply but the downside is that the single parameter does not determine the strength of the spatial structure, and will therefore not accomodate weak correlations and as a result is rather restrictive.

2. **Convolution Model.**
   This model combines intrinsic model but adds a set of random variables to the model and was proposed by Besag (Besag et al., 1991)

   The model is given by;

   $$\phi_k = \phi_k + \psi_k, \phi_k|\delta^2 \sim N(0, \delta^2), \psi = \psi_1, ...\psi_n|W, \tau_1^2 \sim IAR(W, \tau_1^2. \tag{3.9}$$

The downside of this model is that estimating the individual components ϕ_k, ψ_k is not always possible. The MCMC convergence is also slow for this model.

3. **Cressie Model.**
   This model was proposed by Cressie (Cressie, 1992) and Stern H and Cressie (Stern and Cressie, 2000) This model is good because it helps determine the varying strengths of spatial correlations. This model uses a single set of random effects but introduces additional spatial correlation parameters. The model is given by;

   $$\phi_k|\phi_{-k}, W, \rho, \tau^2, \mu \sim N\left(\rho \times \frac{1}{n_k}\sum_{j\sim k}\phi_j + (1-\rho)\mu, \frac{\tau^2}{n_k}\right). \tag{3.10}$$

   This CAR prior has the same conditional variance as the intrinsic model, while the conditional expectation is a weighted average of the mean of the random effects in neighbouring areas and an overall mean μ.
   Here;the weight parameter ρ controls the strength of the spatial correlation between the random effects,with ρ=0 corresponding to independence, while increasing its value towards one corresponds to increasingly strong spatial correlation (ρ=1 simplifies to the intrinsic model).
   The downside of this model is the form of the conditional variance, which is unappealing when ρ is close to zero. This is because in the absence of spatial correlation (when ρ=0) there is no reason for the conditional variance of ϕ_k to be inversely proportional to the number of neighbours, as they provide no information about ϕ_k

4. **Leroux Model.**
   This model was introduced by Leroux (Leroux et al., 2000) and further explained by Madnab (MacNab, 2003) (MacNab et al., 2006). This model is based on a single set of random effects $\phi = (\phi_1, ..., \phi_n)$ which are represented by the multivariate Gaussian distribution.
   The conditional expectation is a weighted average of the random effects in neighbouring areas and the overall mean μ, while the conditional variance has a more attractive form than that in Cressie Model.
   This model is better compared to Cressie because there is no longer any information about ϕ_k in the neighbouring random effects.

3.8 Summary Statistics

In this approach unlike the Bayesian approach where we have CAR priors, we will make use of the data only and thereafter map it to find the results.

We analysed the data and came up with an additional variable on "persons who are non suppressed" and later used the additional variable to map the data to the counties which were used as the reference.

Chloropleth (maps that are shaded with different color intensities)were then used to show the VL suppression spread across all the counties in the Country.

Chapter Four

Results and Findings

4.1 HIV Viral Load Cases

The 2012 to 2017 data on viral load suppression increased gradually over time from 116 centres to 2122 centres in 2017 submitting information on viral load. This means that the number of people also being tested increased over time.

Samples sent in, that were virally suppressed also increased over time significantly as shown below, as well as the number of tests that were taken over the years. The number of samples taken do not necessarily equal the number of people who visited the clinical facilities as one individual may have walked into the facility many times in that particular year or may have visited another clinic over the years. Suppression trends and outcomes are illustrated below as shown on NASCOP website till to date. The graphs indicate that more people are virally suppressed with time as compared to earlier years.

Figure 4.1: Figure showing suppression trends:2012-2019 Source: NASCOP

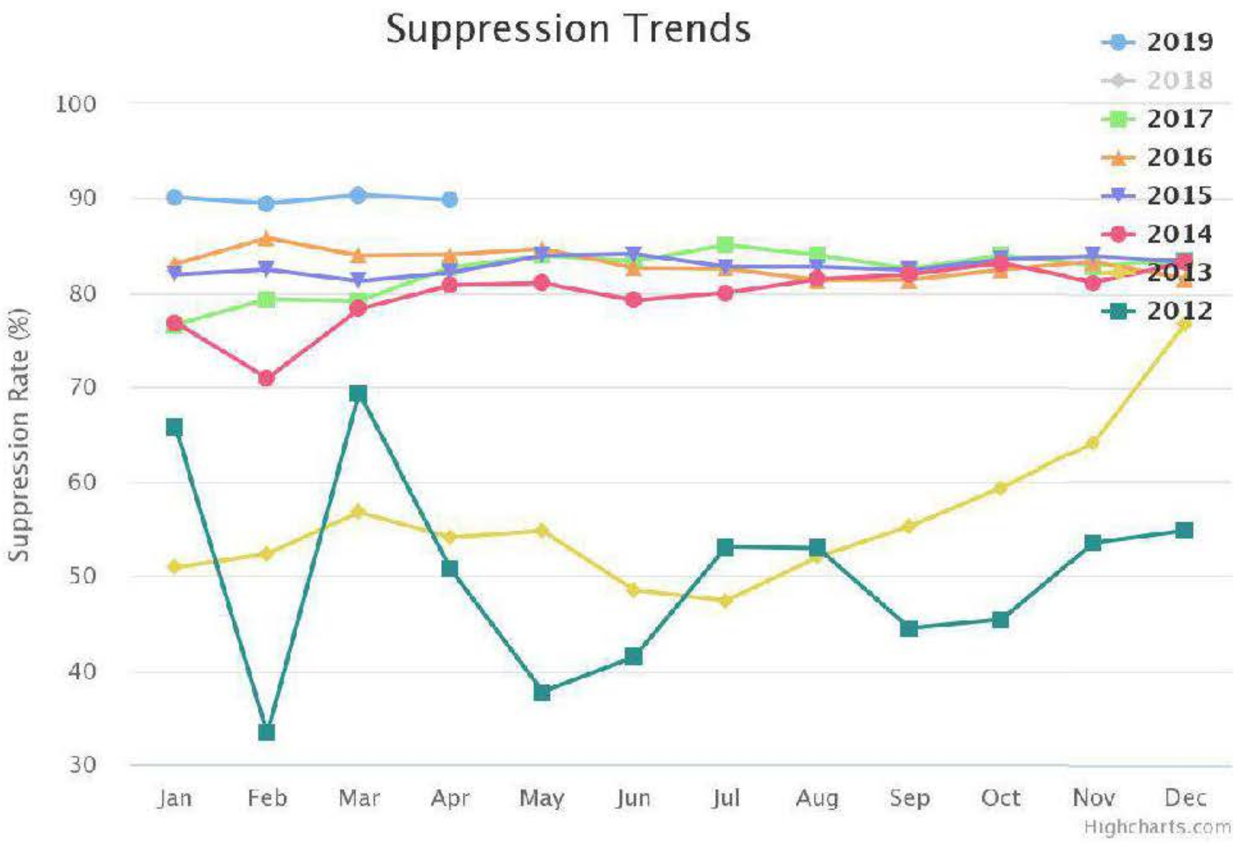
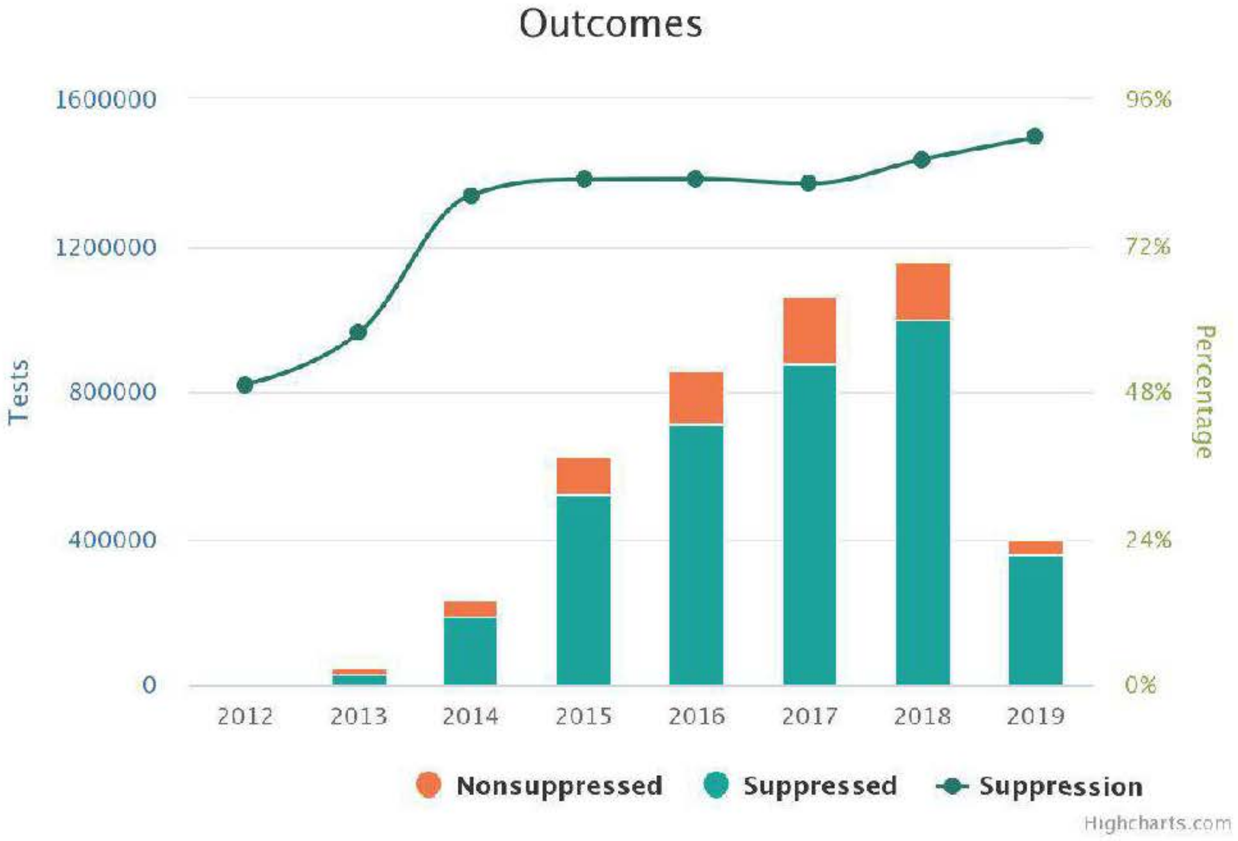


Figure 4.2: Figure showing Suppression Outcomes:2012-2019 Source: NASCOP



The table below is a summary of the variables used, inform of their mean,lower and upper bound means.

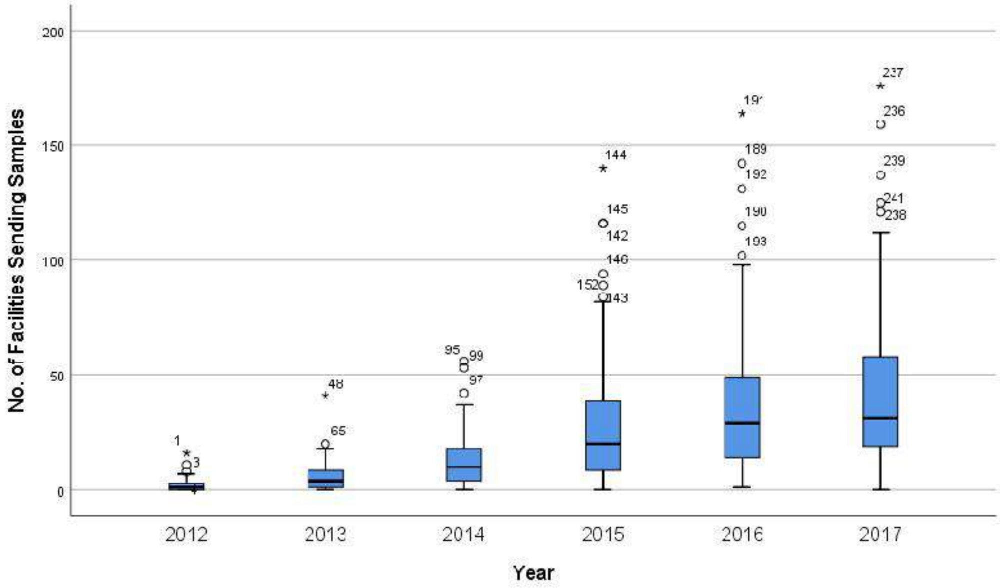
Table 4.1: Table Showing a Summary Description of the Variables in the Data used in the study

Variable Name	Descriptive	2012	2013	2014	2015	2016	2017
No.of.Facilities	Mean	2.47	6.38	14.83	31.70	39.74	45.15
	Lower bound	1.47	8.61	17.29	41.58	51.31	57.69
	Upper Bound	3.46	8.61	19.05	41.58	51.31	57.69
Rejected Samples	Mean	3.21	20.68	39.88	81.64	91.11	156.04
	Lower bound	1.55	10.41	26.52	52.71	54.78	59.16
	Upper Bound	4.87	30.95	53.25	110.56	127.43	252.93
Tests Redrawn	Mean	4.34	22.74	129.36	299.45	520.45	325.13
	Lower bound	1.32	8.95	65.29	113.77	279.63	154.46
	Upper Bound	7.36	36.54	193.42	485.12	761.26	495.80
Tests Taken	Mean	872.87	3115.11	19243.19	51565.15	72775.45	88253.38
	Lower bound	475.98	1817.39	10971.73	30807.01	42026.41	52789.86
	Upper Bound	1269.76	4412.82	27514.65	72323.29	103524.49	123716.90
Perc.VL non.supp	Mean	43.05	46.48	25.32	19.99	19.69	20.09
	Lower bound	36.07	41.49	22.69	18.33	18.01	18.19
	Upper Bound	50.03	51.48	27.93	21.66	21.38	21.98
1st Line Patients	Mean	727	727	2818.67	14910.4	43207.12	62888.72
	Upper bound	392	392	1721.39	8080.49	25858.02	365535.44
	Lower Bound	1062	1062	3915.94	21740.32	60556.24	89242.01
HIV-TB Coinfection	Mean	261.26	261.26	826.69	4081.62	11233.85	16989.15
	Lower bound	152.36	152.36	511.32	2263.96	6723.08	10074.18
	Upper Bound	370.15	370.15	1142.06	5899.27	15744.62	23904.12
Retention Perc	Mean	46.61	46.61	75.42	71.54	75.45	90.13
	Lower bound	39.07	39.07	74.65	68.45	73.25	89.25
	Upper Bound	54.15	54.15	76.19	74.62	77.66	91.00
VL>1000	Mean	431.79	1357.09	3841.12	8912.30	12594.60	14938.34
	Lower bound	253.4	883.16	2453.70	5580.29	7552.06	9357.66
	Upper Bound	610.17	1831.01	5228.54	12244.31	17637.13	20519.02

Data visualization was done through use of box plots. Box plots are ideal in showing the median as well as the variability of the data, in this case the variables, over the period of time that they were collected. The box plot below shows the number of facilities that sent VL samples over between 2012 and 2017, which showed an increase of the same. 2012, few facilities may have been equipped as compared to 2017 hence the steady increase.

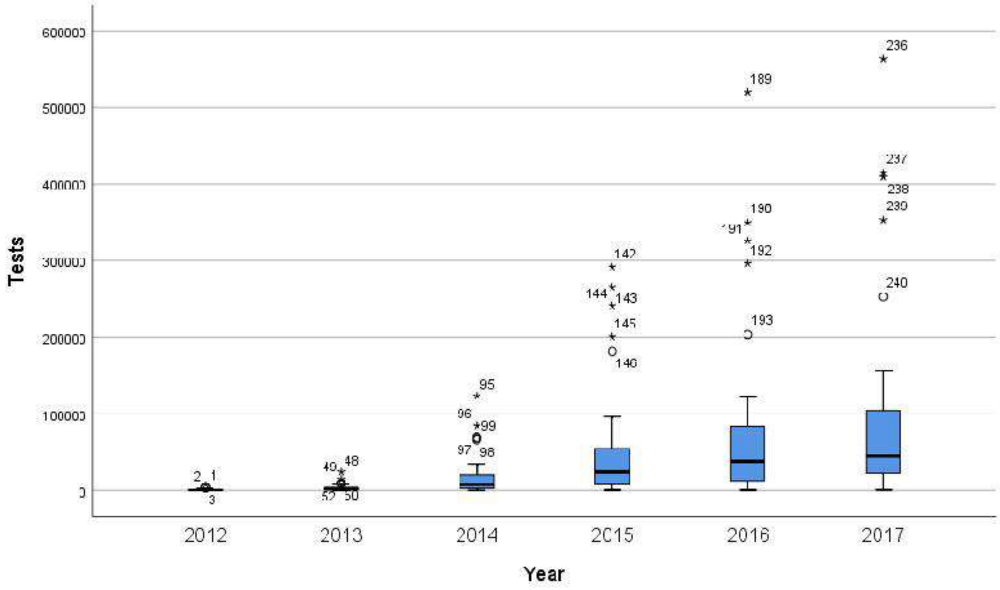
Further, 2015-2017 numbers had more outliers, indicating that the data, showing the general variability of the data.

Figure 4.3: Box Plot showing the number of facilities that sent samples



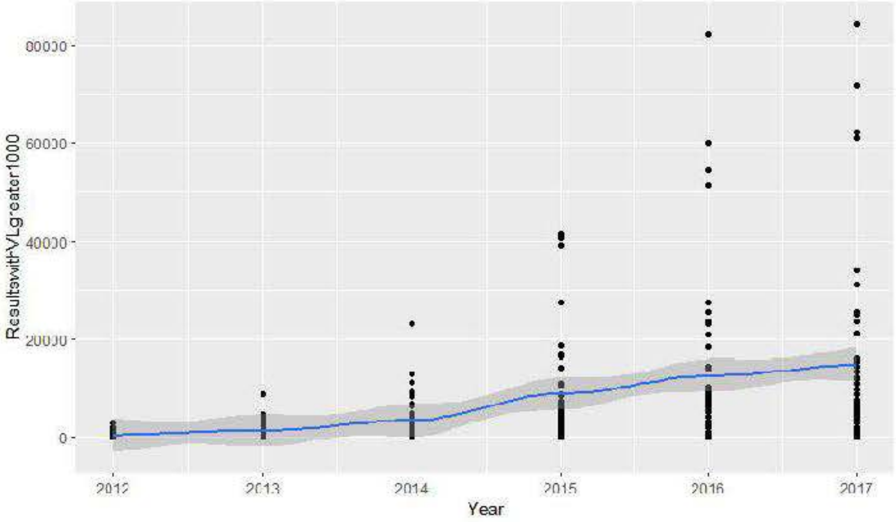
The below box plot shows the median for the tests taken in the centers. Again, there was a steady increase over the period, owing to more awareness and better equipment in the facilities.

Figure 4.4: Box Plot showing the number of samples that were taken in the centers



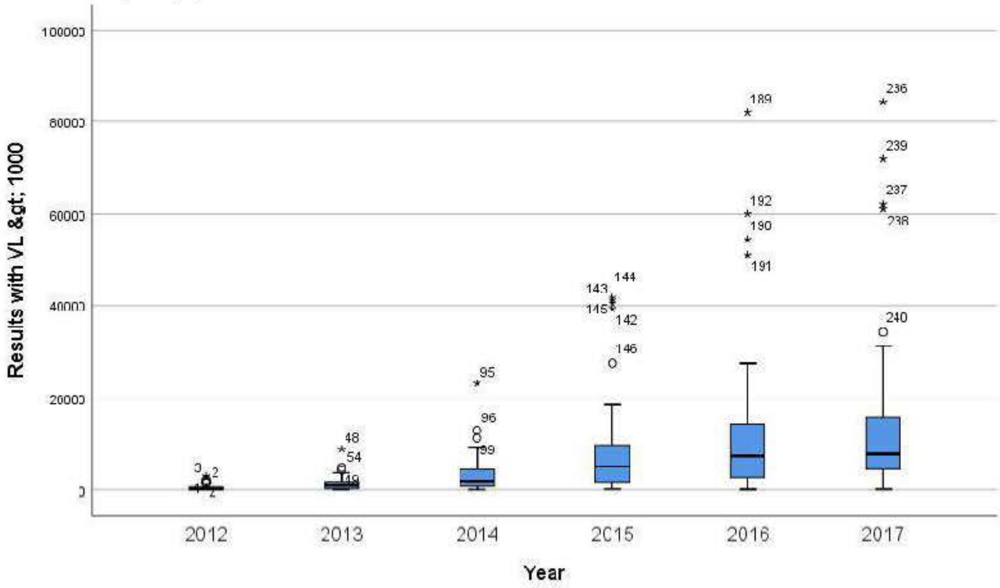
The below trendline shows the number of samples that were non-suppressed over the period. These are samples that contained VL of more than 1000 copies per ml of blood.The same also increased over the period, however, the number of samples taken also increased over the period.

Figure 4.5: Trend Line that shows the number of samples that were not virally suppressed(VL greater than 1000copies/ml)



The box plot below shows the number of samples that were suppressed. These were samples which had VL with less than 1000copies/ml of blood. The box plot shows that the same increased steadily over the period. However, since the number of tests taken also increased over the period,further study in the numbers was needful.

Figure 4.6: Box Plot showing the number of samples that were virally suppressed



4.2 Results: Viral Load Suppression

4.2.1 Summary Statistics Findings

By examining the data at county level, and plotting non-suppression rates, this is what we found. In 2012, the counties that had highest number of virally non-suppressed persons were, Nyandarua, Kakamega, Vihiga, Kajiado and Transzoia Counties. These counties also had the highest retention rates in that same year. However, in terms in HIV-TB co-infection, Nairobi, Kericho, Machakos,

Kisumu, Homa Bay had highest numbers of infected individuals as well as highest numbers of Patients on 1st Line were from these counties. Further, very few facilities were sending samples on Viral Load across the country,which were only 116. This is because Viral Load measurement had just been introduced as initially CD4 count was what was being used. In that same year, counties such as Lamu, West Pokot, Mandera, Tana River and Isiolo did not have any facility that monitored viral load to patients and therefore there were no records of any virally suppressed samples or any samples whatsoever.

In 2013, there were some changes in counties which were least suppressed with Nyamira being at the top as well as Meru, Vihiga, Kakamega, Isiolo while counties such as Kisumu and Busia had high rates of viral suppression. However, Mandera, Wajir and Samburu still had no entries on VL. Most patients on 1st Line were from Kiambu, Kisumu, Uasin Gishu, Busia and Nairobi and the same counties topped in the number of patients with HIV-TB co-infection. In terms of retention, counties such as West Pokot, Uasin Gishu, Mombasa, Tharaka Nithi and Kisii had the most numbers in terms of retention. There was also an increase in facilities that measured the VL from 116 in 2012 to 300 centres in 2013. This must have been due to awareness created steadily over the period.Donor funding is said to have increased as well in that same period from 18.85 per cent in 2012 to 19.6 per cent in 2013.(Kates et al., 2014)

In 2014, 626 centers sent samples on VL, while counties such as Turkana, Mandera, Tharaka Nithi, Wajir and Samburu were the least virally suppressed as compared to Kiambu, Vihiga, Busia, Migori and Kwale whose suppression rates were high. Facilities sending samples increased as well, owing to more awareness and availability of funds in these areas. Lamu, however, did not send any samples on VL that year as well, as opposed to the previous year, 2013.This inconsistency may have been due to little awareness on the importance of VL monitoring or even due to poor facilities.

In 2015,we observed that the facilities sending the samples grew to 1490, as compared to 626 centres in 2014. Counties such as Baringo, Wajir, Samburu, Turkana and Mandera were least suppressed. Populations in these counties is low as compared to other counties, and may therefore not necessarily mean that these counties were doing poorly. However,counties such as Kiambu, Nairobi, Nyeri, Kirinyaga and Meru had high number of suppression rates, meaning that most samples sent had VL of less than 1000copies/ml of blood. Further,Migori, Siaya, Homabay, Kisumu and Nairobi counties had the highest number of individuals with HIV-TB co-infections as well as patients on 1st Line Treatment. High retention rates were observed in Meru, Kirinyaga, Nyeri,Nairobi and Kiambu counties. We also observe that these counties are among the ones whose population is suppressed.

In 2016, the facilities that were sending samples on VL rose to 1868, as compared to 1426 in 2015. Retention rates in Nyandarua, Kiambu, Migori, Meru and Kirinyaga were highest.This means that most of the patients in these counties continued in the treatment in that particular year. Population in Counties such as Turkana,Samburu, Mandera,Tana River,Elgeyo Marakwet were most suppressed as opposed to people in Kirinyaga, Meru, Migori, Kiambu and Nyandarua counties. Lastly, Nairobi, Kisumu,

Homa Bay, Siaya and Migori had the highest number of persons on 1st Line treatment as well as HIV-TB Co-infection.

In 2017,the total number of facilities that sent the samples was 2122.This was approximately 13 per cent increase from 2016 and over 170 per cent from 2012. Nyeri, Kiambu, Kirinyaga, Muranga, Kisii had highest retention rates as compared to Elgeyo Marakwet Mandera, Tana River, Samburu, Turkana which had least retention rates.Nairobi, Homa Bay, Kisumu Siaya and Migori had highest number of people with HIV-TB co-infection as well as patients on 1st Line treatment. In terms of suppression, Samburu, Turkana, Tana River, Mandera and Elgeyo Marakwet counties are least suppressed.

High retention did not necessarily mean high suppression rates as will be clearer in the Bayesian Approach.
Below maps show the regions that are least virally suppressed; The areas that had deeper color meant that the people in these regions were highly virally non-suppressed.

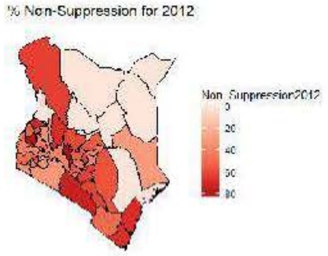


Figure 4.7: Map of nonsuppression 2012

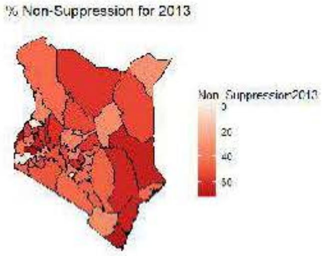


Figure 4.8: Map of nonsuppression 2013

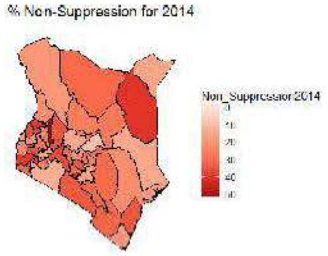


Figure 4.9: Map of nonsuppression 2014

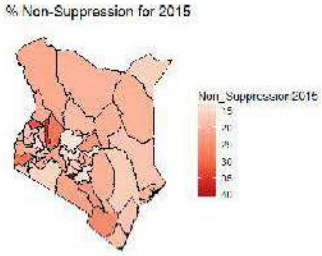


Figure 4.10: Map of nonsuppression 2015

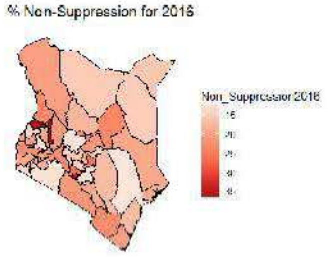


Figure 4.11: Map of nonsuppression 2016

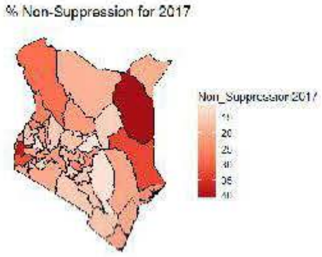


Figure 4.12: Map of nonsuppression 2017

4.2.2 Bayesian Approach

Under Bayesian paradigm, we employed spatial models in order to provide insight of counties high non-suppression. This provided a further analysis of predictors of non-suppression.

Spatial autocorrelation was also measured to assess the extent of the existence of autocorrelation in the data using the global Moran's I statistic. The below table shows that the data was spatially autocorrelated owing from the small p values except in 2012, where the p-value was high.Small p values show evidence against the null hypodissertation which in this case is that there exists no autocorrelation in the data.Existence of spatial autocorrelation meant that counties close to each other had almost similar outcomes as compared to those further from each other. Below is a summary of the findings;

Table 4.2: Table Showing a Summary of the Moran's Index from 2012 to 2017

Year	Moran's I standard deviate	Moran's I statistics	P value	Expectation	Mean
2012	0.071422	0.8784906	0.2143	-0.0149	-0.121
2013	0.081812	0.7062876	0.027	-0.0149	-0.0137
2014	0.072202	0.8880846	0.0397	-0.0149	-0.0139
2015	0.082592	0.6158816	0.0524	-0.0149	-0.0158
2016	0.072982	0.8976786	0.0451	-0.0149	-0.0152
2017	0.093372	0.7484756	0.0778	-0.0149	-0.0337

To determine the association of VL suppression and determinants, a Bayesian approach was applied where Conditional Autoregressive Model (CAR) was specified. This model depended on the conditional distribution of spatial error terms and explained part of the variability of the relative risk.

Prior to modelling, we specified the model using the BUGS (Bayesian inference using Gibbs sampler) language, the VL non-suppression data, spatial data describing the neighbourhood structure and initial values of the parameters. At analysis stage, MCMC simulation method was used and 100,000 iterations specified where the first 10,000 were discarded leaving 90000 and each 100th sample was stored. The covariates were included in order to assess and remove the effects that occured as a result of confounding factors.

Further, in Bayesian approach, we modelled expected VL using retention rates, and HIV-TB co-infections as priors to come up with expected VL.

In 2012 it shows that counties such as Nyandarua and Kisumu were least suppressed whereas counties such as Wajir, West Pokot and Garissa were highly suppressed, but this changes over time.

In 2013 however, we note Homabay is among counties with high viral suppression as well as Nandi Counties,however,Busia, Kwale and Kisumu are least virally suppressed.This might be due to the fact that they are counties that border Tanzania and Uganda, hence translating to alot of movement across the 3 countries leading to little retention and therefore low viral suppression.

In 2014 this changes. Migori is among the counties that are least virally suppressed while other counties such as Mandera,Wajir, Garissa and Turkana follow closely after Migori. Counties such as Isiolo, Kajiado and Makueni are seen to be doing well in terms of suppression.

In 2015,most counties are highly suppressed.This might be because of awareness now created.However, Nairobi, is still highly non-suppressed still,probably due to the high population in the area. 2016, this changes, as Homabay tops up the list of counties that are non-suppressed with Garissa and Isiolo following closely behind.

Lastly, in 2017, still Homabay, Garissa and Isiolo are still among top counties that are least virally suppressed as well as Kirinyaga, while counties such as kajiado,Kilifi and Lamu are doing much better in terms of VL suppression. Below maps show the regions that are least virally suppressed;

The areas that have deeper color meant that the people in these regions were highly virally suppressed.

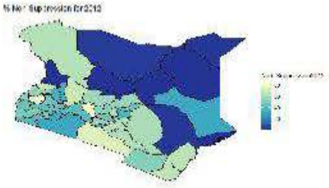


Figure 4.13: Bayesian:non-suppression 2012

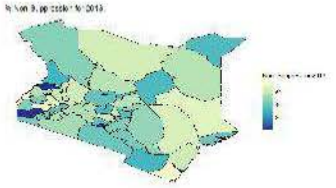


Figure 4.14: Bayesian:non-suppression 2013

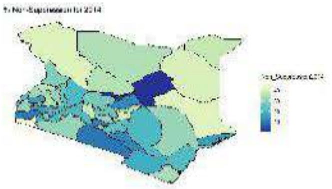


Figure 4.15: Bayesian:non-suppression 2014

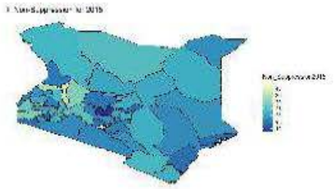


Figure 4.16: Bayesian:non-suppression 2015

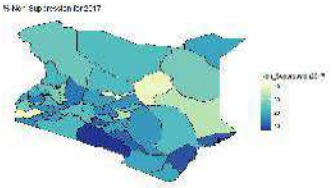


Figure 4.17: Bayesian:non-suppression 2016

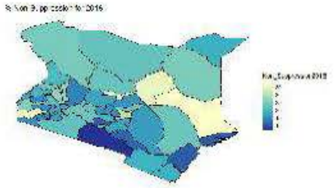


Figure 4.18: Bayesian:non-suppression 2017

4.3 Discussion

In this study we have assessed the spatial distribution of VL suppression in Kenya at different counties from 2012 to 2017.
In addition we have identified areas with high VL non-suppression using Bayesian smoothed maps, Global Moran's I statistics and the traditional summary statistics approach.
We also modelled the covariates that had an influence on VL non-suppression. For exploratory spatial data analysis, Bayesian smoothed maps of standardized non-VL suppression were generated.

Standardized non-VL suppression was modelled to obtain accurate rates, and for comparison. This is because, areas with lower population might have indicated high suppression as compared to high population areas. This is one weakness of the traditional summary statistics approach we used that Bayesian approach covers for.

This allowed visualization of spatial VL suppression patterns in 47 counties in Kenya. Results showed that Homabay, Isiolo and Garissa were least virally suppressed in 2017. In addition most of the urban areas have recorded high viral suppression rates, such as Nairobi. Homabay is still highly non-suppressed for 2 consecutive years. Homabay's low suppression rates is attributed to the fishing that goes on in the area. Mostly, residents in the area engage in sexual activities in exchange for fish, a popular delicacy in that area. Other factors may be due to movement of residents from Kenya moving over to Uganda and Tanzania, which may attribute to low retention and therefore adherence is poor.Wife inheritance is also a popular culture that has been in place since independence.

Although the practice is getting unpopular, this may be one of the reasons for the low suppression rates. Usually, multiple sex partners can contribute to a new HIV strain that may require a different HIV regimen altogether. Isiolo and Garissa are popular with the Borana, and Somali and Meru people. These groups practice polygamy however, this might not be necessarily a reason for high HIV non-suppression rates. Major reason is attributed to literacy levels being low,currently rated at 8 per cent by a Kenya National Adult Literacy Survey report that also indicated Nairobi County had the highest literacy levels, leading to low and high awareness respectively.

Global Moran's I for spatial autocorrelation computed showed that counties closer to each other had similar relative VL non-suppression as compared to those further away. This relationship is significant in 2013 to 2017. We were also able to detect areas of decreasing or increasing trends of VL suppression in relationship to their neighbours.

The Bayesian regression was also fitted to determine predictors of VL Non-suppression. from the results the 3 factors HIV/TB co-infection, 1st Line Regimens and retention rates were significant. kld is interpreted as p-value. The model has no random effects, no hyperparameters. The expected number of effective parameters was (stddev): 13.46(0.00), with number of equivalent replicates 113.30, deviance Information Criterion: 354.48, effective number of parameters: 4.973 and marginal likelihood:-336.0. see table below

Fixed Effects	Mean	SD	0.025 quantile	0.5 quantile	0.975 quantile	kld
Intercept	5.12	1.11	2.89	6.44	9.32	0.00011094
HIV/TB coinfection	21.70	1.55	10.32	23.02	33.34	0.00000980
Ist Line Regimen	23069.04	2133.00	10345.45	23070.36	33415.81	0.00001375
Retention Rate	46.16	2.54	21.29	47.48	68.78	0.00000746

Chapter Five

Conclusion and Recommendations

This chapter entails a conclusion on the findings of this dissertation as well as recommendations for further research.

5.1 Conclusion

Bayesian Models are good because Bayesian inference allows informative priors so that prior knowledge or results of a previous model can be used to inform the current model.Also, Bayesian inference uses prior distributions, so more information is used and 95 per cent probability intervals of posterior distributions should be narrower than 95 per cent confidence intervals of point-estimates.

Traditional models (whereby non-standardized data is used) are usually easier to prepare because many things do not need to be specified, such as prior distributions, initial values for numerical approximation,and usually the likelihood function unlike in Bayesian approach.

In this study we found retention as a significant factor and positively affected VL suppression. Patients who were retained for a long time were likely to have their VL suppressed.Although measurement of retention rates has been deemed to be complex, as it involves multiple visits over a period of time (Mugavero et al., 2012).

1st Line ART Regimen and HIV-TB Co-infection go hand in hand as patients who are on 1st Line ART regimen are most likely to have TB as well. However, in this study, we identified that these two factors were significant variables in determining VL suppression.Co-infections contribute to HIV-related pathogenesis and often increase viral load in HIV-infected people (Modjarrad and Vermund, 2010), however this is not the case always as HIV patients who have TB but are on Anti-TB drugs have been found to have VL below 1000copies/ml once they started treatment of TB.

5.2 Recommendations for further research

The work focused only on county level data with the specified variables used to interpret the outcome on viral suppression. Spatio-temporal methods could also be used to examine VL suppression data in future studies. Also, viral load suppression can be examined among different demographic groups, as in this research we only focused on county level data only without breaking it down into various demographic populations.

Bibliography

Besag, J. and Higdon, D. (1999). Bayesian analysis of agricultural field experiments. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61(4):691–746.

Besag, J., York, J., and Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the institute of statistical mathematics*, 43(1):1–20.

Boerma R, B. T. and J, C. (2016). Suboptimal viral suppression rates among hiv-infected children in low-and middle-income countries: a meta-analysis. *Clinical Infectious Diseases*, pages 1645–1654.

Braitstein, P., Katshcke, A., Shen, C., Sang, E., Nyandiko, W., Ochieng, V. O., Vreeman, R., Yiannoutsos, C. T., Wools-Kaloustian, K., and Ayaya, S. (2010). Retention of hiv-infected and hiv-exposed children in a comprehensive hiv clinical care programme in western kenya. *Tropical Medicine & International Health*, 15(7):833–841.

Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American statistical Association*, 88(421):9–25.

Casetti, E. and Fan, C. (1991). The spatial spread of the aids epidemic in ohio: empirical analyses using the expansion method. *Environment and Planning A*, 23(11):1589–1608.

CDC (2005). *Compendium of measures to prevent disease associated with animals in public settings, 2005*, volume 54. US Dept. of Health and Human Services.

Cheluget, B., Baltazar, G., Orege, P., Ibrahim, M., Marum, L. H., and Stover, J. (2006). Evidence for population level declines in adult hiv prevalence in kenya. *Sexually transmitted infections*, 82(suppl 1):i21–i26.

Clarke, K. C., McLafferty, S. L., and Tempalski, B. J. (1996). On epidemiology and geographic information systems: a review and discussion of future directions. *Emerging infectious diseases*, 2(2):85.

Cohen, M. S., Chen, Y. Q., McCauley, M., Gamble, T., Hosseinipour, M. C., Kumarasamy, N., Hakim, J. G., Kumwenda, J., Grinsztejn, B., Pilotto, J. H., et al. (2011). Prevention of hiv-1 infection with early antiretroviral therapy. *New England journal of medicine*, 365(6):493–505.

Cressie, N. (1992). Statistics for spatial data. *Terra Nova*, 4(5):613–617.

Doshi, R. K., Milberg, J., Jumento, T., Matthews, T., Dempsey, A., and Cheever, L. W. (2017). For many served by the ryan white hiv/aids program, disparities in viral suppression decreased, 2010–14. *Health Affairs*, 36(1):116–123.

Gardner, E. M., McLees, M. P., Steiner, J. F., del Rio, C., and Burman, W. J. (2011). The spectrum of engagement in hiv care and its relevance to test-and-treat strategies for prevention of hiv infection. *Clinical infectious diseases*, 52(6):793–800.

Geanuracos, C. G., Cunningham, S. D., Weiss, G., Forte, D., Henry Reid, L. M., and Ellen, J. M. (2007). Use of geographic information systems for planning hiv prevention interventions for high-risk youths. *American Journal of Public Health*, 97(11):1974–1981.

Ickovics, J. R. and Meade, C. (2002). Adherence to haart among patients with hiv: breakthroughs and barriers. *AIDS care*, 14(3):309–318.

Kandwal, R., Garg, P., and Garg, R. (2009). Health gis and hiv/aids studies: Perspective and retrospective. *Journal of biomedical informatics*, 42(4):748–755.

Kates, J., Wexler, A., and Lief, E. (2014). Financing the response to hiv in low and middle-income countries. *Menlo Park, CA: Kaiser Family Foundation*.

Leroux, B. G., Lei, X., and Breslow, N. (2000). Estimation of disease rates in small areas: a new mixed model for spatial dependence. In *Statistical models in epidemiology, the environment, and clinical trials*, pages 179–191. Springer.

Levi, J., Raymond, A., Pozniak, A., Vernazza, P., Kohler, P., and Hill, A. (2016). Can the unaids 90-90-90 target be achieved? a systematic analysis of national hiv treatment cascades. *BMJ global health*, 1(2):e000010.

Li, H., Calder, C. A., and Cressie, N. (2007). Beyond moran's i: testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis*, 39(4):357–375.

MacNab, Y. C. (2003). Hierarchical bayesian modeling of spatially correlated health service outcome and utilization rates. *Biometrics*, 59(2):305–315.

MacNab, Y. C., Kmetic, A., Gustafson, P., and Sheps, S. (2006). An innovative application of bayesian disease mapping methods to patient safety research: A canadian adverse medical event study. *Statistics in medicine*, 25(23):3960–3980.

Modjarrad, K. and Vermund, S. H. (2010). Effect of treating co-infections on hiv-1 viral load: a systematic review. *The Lancet infectious diseases*, 10(7):455–463.

Moran, P. (1950). Notes on continuous stochastic phenomena biometrika 37: 17–23. *Find this article online*.

Mugavero, M. J., Westfall, A. O., Zinski, A., Davila, J., Drainoni, M.-L., Gardner, L. I., Keruly, J. C., Malitz, F., Marks, G., Metsch, L., et al. (2012). Measuring retention in hiv care: the elusive gold standard. *Journal of acquired immune deficiency syndromes (1999)*, 61(5):574.

NACC (2016). Kenya aids response progress report 2016.

Nicole, L., Mao, L., De Wit, J., Smith, D., Holt, M., Prestage, G., Wilson, D. P., and Petoumenos, K. (2015). Estimating antiretroviral treatment coverage rates and viral suppression rates for homosexual men in australia. *Sexual health*, 12(5):453–457.

Onywera, H., Maman, D., Inzaule, S., Auma, E., Were, K., Fredrick, H., Owiti, P., Opollo, V., Etard, J.-F., Mukui, I., et al. (2017). Surveillance of hiv-1 pol transmitted drug resistance in acutely and recently infected antiretroviral drug-naive persons in rural western kenya. *PloS one*, 12(2):e0171124.

Sidibé, M., Loures, L., and Samb, B. (2016). The unaids 90–90–90 target: a clear choice for ending aids and for sustainable health and development. *Journal of the International AIDS Society*, 19(1).

Stern, H. S. and Cressie, N. (2000). Posterior predictive model checks for disease mapping models. *Statistics in medicine*, 19(17-18):2377–2397.

Tobias, C., Cunningham, W. E., Cunningham, C. O., and Pounds, M. B. (2007). Making the connection: the importance of engagement and retention in hiv medical care. *AIDS patient care and STDs*, 21(S1):S–3.

UNAIDS (2017). Latest statistics on the status of the aids epidemic.

Waller, L. A. and Gotway, C. A. (2004). *Applied spatial statistics for public health data*, volume 368. John Wiley & Sons.

Appendices

R scripts

```r
library(maptools)
library(raster)
library(plyr)
library(dplyr)
library(ggplot2)
library(rgdal)
gpclibPermit()
Kenya<-getData("GADM", country="KE", level=0)
Kenya1<-getData("GADM", country="KE", level=1)

Kenya1_UTM<-spTransform(Kenya1, CRS("+init=EPSG:32737"))

NAME_1<-Kenya1_UTM@data$NAME_1
Non_Suppression2012      <-      c(66.7,33.8,53.2,49.1,33.3,36.3,29.6,49.5,
0.0,72.9,79.8,40.9,54.8,67.5,60.3,43.6,37.2,
56.6,66.6,59.5,0.0,61.7,60.7,0.0,0.0,54.4,
35.9,63.3,43.0,37.9,38.0,42.2,31.1,53.2,83.5,
52.8,0.0,40.8,49.3,0.0,42.0,70.2,62.2,32.5,77.4,0.0,0.0)
Non_Suppression2013      <-      c(59.4,44.7,41.6,31.4,39.1,65.3,59.5,45.1,
73.9,46.7,69.9,52.6,35.8,44.0,42.9,47.5,30.7,
58.4,57.9,47.1,33.3,61.6,55.7,0.0,51.1,65.9,50.0,
47.3,53.6,35.8,54.2,40.6,32.6,74.5,64.0,47.5,0.0,
44.3,57.7,38.8,53.8,34.6,62.1,33.4,66.2,0.0,32.4)
Non_Suppression2014      <-      c(33.7,32.9,19.0,14.0,30.3,33.7,27.6,
16.4,34.8,29.5,18.9,24.9,15.0,17.8,17.2,
29.5,15.2,30.3,8.7,30.3,0.0,32.9,35.7,40.7,
35.2,26.0,13.2,35.3,28.9,18.8,25.5,21.5,30.9,
18.3,21.2,28.3,38.3,19.1,24.8,27.0,39.4,21.5,
50.5,23.9,14.8,38.7,26.4)
Non_Suppression2015      <-      c(27.8,21.7,18.8,17.0,16.2,16.9,19.6,16.8,19.1,
19.6,17.8,20.3,12.9,18.9,
13.6,22.9,15.5,19.4,23.1,
16.5,18.6,16.9,17.1,41.4,21.0,
13.8,15.2,18.7,14.5,
13.5,18.5,21.4,25.2,18.7,16.0,
13.6,31.9,20.7,19.3,
27.2,19.5,20.1,35.4,21.2,17.2,
28.6,19.9)
Non_Suppression2016      <-      c(20.9,21.2,20.8,20.9,
26.1,16.7,18.6,15.7,22.3,
16.8,20.3,17.8,12.3,18.7,
12.9,18.8,15.6,15.2,21.3,
13.6,22.8,14.7,14.0,34.3,
22.5,12.6,12.6,19.3,14.6,
15.8,17.7,22.4,22.6,20.8,
12.2,14.3,35.6,20.3,19.3,
29.3,15.0,23.1,36.6,25.5,19.2,19.1,23.3)
Non_Suppression2017      <-      c(15.0,15.0,14.9,20.3,13.5,20.0,17.7,15.3,
```

```
19.5,11.7,20.7,13.1,16.1,23.5,15.8,16.2,
19.0,18.4,11.2,19.5,11.8,23.7,16.3,19.6,
21.5,11.7,19.6,18.8,16.9,19.0,21.6,38.6,
15.8,21.9,20.1,22.6,24.6,28.2,20.0,26.6,
41.1,25.3,19.9,23.8,32.8,31.7,14.4)

Non_Suppression2012_df<-data.frame(NAME_1, Non_Suppression2012)

Kenya1_UTM@data$id <- rownames(Kenya1_UTM@data)
Kenya1_UTM@data <- join(Kenya1_UTM@data, Non_Suppression2012_df,
by="NAME_1")
Kenya1_df <- fortify(Kenya1_UTM)
Kenya1_df <- join(Kenya1_df,Kenya1_UTM@data, by="id")


library(ggmap)
library(scales)

theme_opts<-list(theme(panel.grid.minor = element_blank(),
                       panel.grid.major = element_blank(),
                       panel.background = element_blank(),
                       plot.background = element_blank(),
                       axis.line = element_blank(),
                       axis.text.x = element_blank(),
                       axis.text.y = element_blank(),
                       axis.ticks = element_blank(),
                       axis.title.x = element_blank(),
                       axis.title.y = element_blank(),
                       plot.title = element_blank()))

ggplot() +
  geom_polygon(data = Kenya1_df, aes(x = long, y = lat, group = group,
 fill =


Non_Suppression2012), color = "black", size = 0.25) +
  theme(aspect.ratio=1)+
  scale_fill_distiller(name="Non_Suppression2012", palette = "Reds", trans = "r
  theme_void()+
  labs(title="%␣Non-Suppression␣for␣2012")

Non_Suppression2013_df<-data.frame(NAME_1, Non_Suppression2013)

Kenya1_UTM@data$id <- rownames(Kenya1_UTM@data)
Kenya1_UTM@data <- join(Kenya1_UTM@data, Non_Suppression2013_df,
by="NAME_1")
Kenya1_df <- fortify(Kenya1_UTM)
Kenya1_df <- join(Kenya1_df,Kenya1_UTM@data, by="id")

theme_opts<-list(theme(panel.grid.minor = element_blank(),
```

```
                          panel.grid.major = element_blank(),
                          panel.background = element_blank(),
                          plot.background = element_blank(),
                          axis.line = element_blank(),
                          axis.text.x = element_blank(),
                          axis.text.y = element_blank(),
                          axis.ticks = element_blank(),
                          axis.title.x = element_blank(),
                          axis.title.y = element_blank(),
                          plot.title = element_blank()))

ggplot() +
  geom_polygon(data = Kenya1_df, aes(x = long, y = lat, group = group, fill =

Non_Suppression2013), color = "black", size = 0.25) +
  theme(aspect.ratio=1)+
  scale_fill_distiller(name="Non_Suppression2013", palette = "Reds", trans = "r
  theme_void()+
  labs(title="%␣Non-Suppression␣for␣2013")

Non_Suppression2014_df<-data.frame(NAME_1, Non_Suppression2014)

Kenya1_UTM@data$id <- rownames(Kenya1_UTM@data)
Kenya1_UTM@data <- join(Kenya1_UTM@data, Non_Suppression2014_df,
by="NAME_1")
Kenya1_df <- fortify(Kenya1_UTM)
Kenya1_df <- join(Kenya1_df,Kenya1_UTM@data, by="id")


theme_opts<-list(theme(panel.grid.minor = element_blank(),
                          panel.grid.major = element_blank(),
                          panel.background = element_blank(),
                          plot.background = element_blank(),
                          axis.line = element_blank(),
                          axis.text.x = element_blank(),
                          axis.text.y = element_blank(),
                          axis.ticks = element_blank(),
                          axis.title.x = element_blank(),
                          axis.title.y = element_blank(),
                          plot.title = element_blank()))

ggplot() +
  geom_polygon(data = Kenya1_df, aes(x = long, y = lat, group = group, fill =
Non_Suppression2014), color = "black", size = 0.25) +
  theme(aspect.ratio=1)+
  scale_fill_distiller(name="Non_Suppression2014", palette = "Reds", trans = "r
breaks = pretty_breaks(n = 5))+
  theme_void()+
  labs(title="%␣Non-Suppression␣for␣2014")
```

```
Non_Suppression2015_df<-data.frame(NAME_1, Non_Suppression2015)

Kenya1_UTM@data$id <- rownames(Kenya1_UTM@data)
Kenya1_UTM@data <- join(Kenya1_UTM@data, Non_Suppression2015_df, by="NAME_1")
Kenya1_df <- fortify(Kenya1_UTM)
Kenya1_df <- join(Kenya1_df,Kenya1_UTM@data, by="id")


theme_opts<-list(theme(panel.grid.minor = element_blank(),
                       panel.grid.major = element_blank(),
                       panel.background = element_blank(),
                       plot.background = element_blank(),
                       axis.line = element_blank(),
                       axis.text.x = element_blank(),
                       axis.text.y = element_blank(),
                       axis.ticks = element_blank(),
                       axis.title.x = element_blank(),
                       axis.title.y = element_blank(),
                       plot.title = element_blank()))

ggplot() +
  geom_polygon(data = Kenya1_df, aes(x = long, y = lat, group = group, fill =
  Non_Suppression2015), color = "black", size = 0.25) +
  theme(aspect.ratio=1)+
  scale_fill_distiller(name="Non_Suppression2015", palette = "Reds", trans = "r
theme_void()+
labs(title="%␣Non-Suppression␣for␣2015")

Non_Suppression2016_df<-data.frame(NAME_1, Non_Suppression2016)

Kenya1_UTM@data$id <- rownames(Kenya1_UTM@data)
Kenya1_UTM@data <- join(Kenya1_UTM@data, Non_Suppression2016_df, by="NAME_1")
Kenya1_df <- fortify(Kenya1_UTM)
Kenya1_df <- join(Kenya1_df,Kenya1_UTM@data, by="id")


theme_opts<-list(theme(panel.grid.minor = element_blank(),
                       panel.grid.major = element_blank(),
                       panel.background = element_blank(),
                       plot.background = element_blank(),
                       axis.line = element_blank(),
                       axis.text.x = element_blank(),
                       axis.text.y = element_blank(),
                       axis.ticks = element_blank(),
                       axis.title.x = element_blank(),
                       axis.title.y = element_blank(),
                       plot.title = element_blank()))

ggplot() +
  geom_polygon(data = Kenya1_df, aes(x = long, y = lat, group = group, fill =
```

```
  Non_Suppression2016), color = "black", size = 0.25) +
  theme(aspect.ratio=1)+
  scale_fill_distiller(name="Non_Suppression2016", palette = "Reds", trans = "r
  theme_void()+
  labs(title="%␣Non-Suppression␣for␣2016")

Non_Suppression2017_df<-data.frame(NAME_1, Non_Suppression2017)

Kenya1_UTM@data$id <- rownames(Kenya1_UTM@data)
Kenya1_UTM@data <- join(Kenya1_UTM@data, Non_Suppression2017_df, by="NAME_1")
Kenya1_df <- fortify(Kenya1_UTM)
Kenya1_df <- join(Kenya1_df,Kenya1_UTM@data, by="id")


theme_opts<-list(theme(panel.grid.minor = element_blank(),
                       panel.grid.major = element_blank(),
                       panel.background = element_blank(),
                       plot.background = element_blank(),
                       axis.line = element_blank(),
                       axis.text.x = element_blank(),
                       axis.text.y = element_blank(),
                       axis.ticks = element_blank(),
                       axis.title.x = element_blank(),
                       axis.title.y = element_blank(),
                       plot.title = element_blank()))

ggplot() +
  geom_polygon(data = Kenya1_df, aes(x = long, y = lat, group = group, fill =
  Non_Suppression2017), color = "black", size = 0.25) +
  theme(aspect.ratio=1)+
  scale_fill_distiller(name="Non_Suppression2017", palette = "Reds", trans = "r
  theme_void()+
  labs(title="%␣Non-Suppression␣for␣2017")
```

Bayesian Results

Table 1: Table Showing Bayesian Anova Results

ANOVAa,b					
Source	Sum of Squares	df	Mean Square	F	Sig.
Regression	54763.759	8	6845.470	65.142	.000
Residual	28162.726	268	105.085		
Total	82926.486	276			
a.Dependent Variable: Percentage VL Non-suppression					
b. Model: (Intercept), Year, Patientson1stLine, HIV-TBCoinfection, Retention-Percent					

Table 2: Table Showing Bayesian Estimates of Coefficients

Bayesian Estimates of Coefficients a,b,c					
	Posterior			95 per cent Credible Interval	
Parameter	Mode	Mean	Variance	Lower Bound	Upper Bound
(Intercept)	-23.768	-23.768	15.153	-31.403	-16.132
Year = 2012	42.668	42.668	8.045	37.105	48.231
Year = 2013	46.104	46.104	8.045	40.540	51.667
Year = 2014	10.081	10.081	5.862	5.332	14.830
Year = 2015	7.331	7.331	5.483	2.738	11.924
Year = 2016	6.230	6.230	5.144	1.781	10.679
Year = 2017	.d	.d	.d	.d	.d
Patientson1stLine	2.433E-6	2.433E-6	.000	.000	.000
HIV-TBCoinfection	.000	.000	.000	.002	.002
Retention-Percent	.519	.519	.002	.442	.596
a. Dependent Variable: Percentage VL Non-suppression					
b. Model: (Intercept), Year, Patientson1stLine, HIV-TBCoinfection, Retention-Percent					
c. Assume standard reference priors.					
d. This parameter is redundant. Posterior statistics are not calculated.					

Table 3: Table Showing Bayesian Estimates of Error Variance

Bayesian Estimates of Error Variance,a					
	Posterior			95 per cent Credible Interval	
Parameter	Mode	Mean	Variance	Lower Bound	Upper Bound
Error variance	104.306	105.875	84.920	89.337	125.420
a. Assume standard reference priors					

Figure 1: Graph showing likelihood functions for patients on 1st Line Regimen

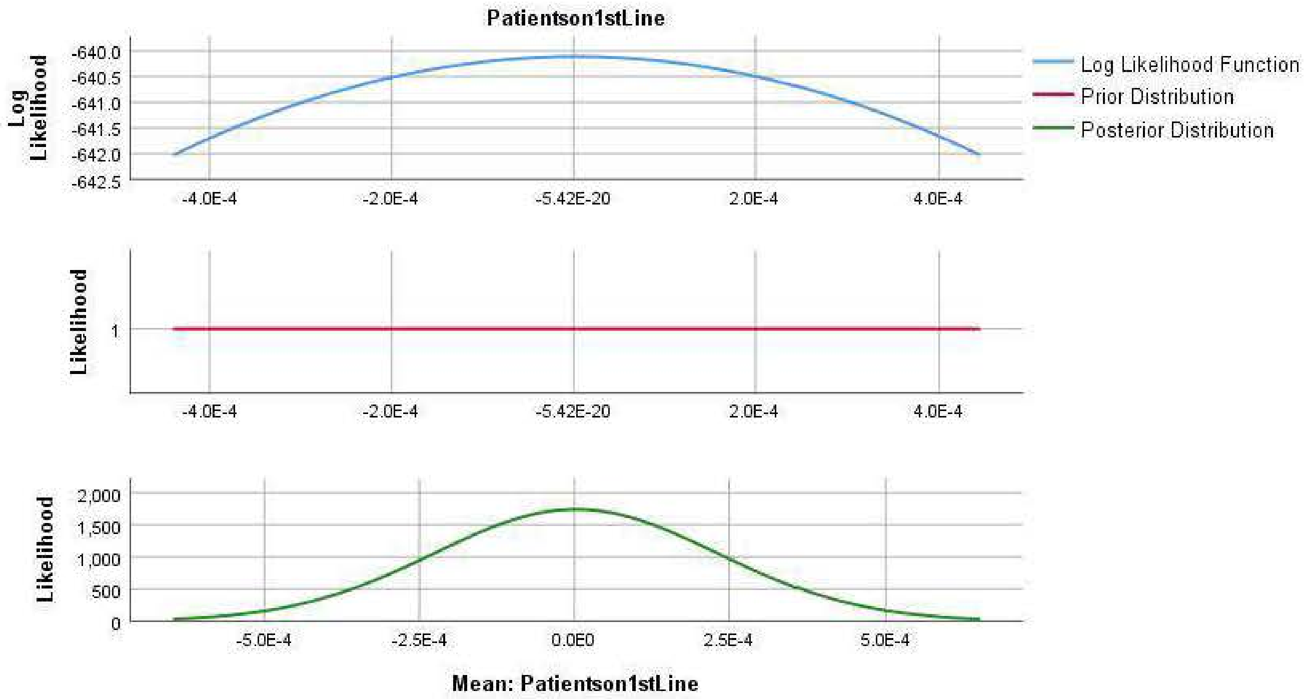


Figure 2: Graph showing likelihood functions for HIV-TB Co-infection Patients

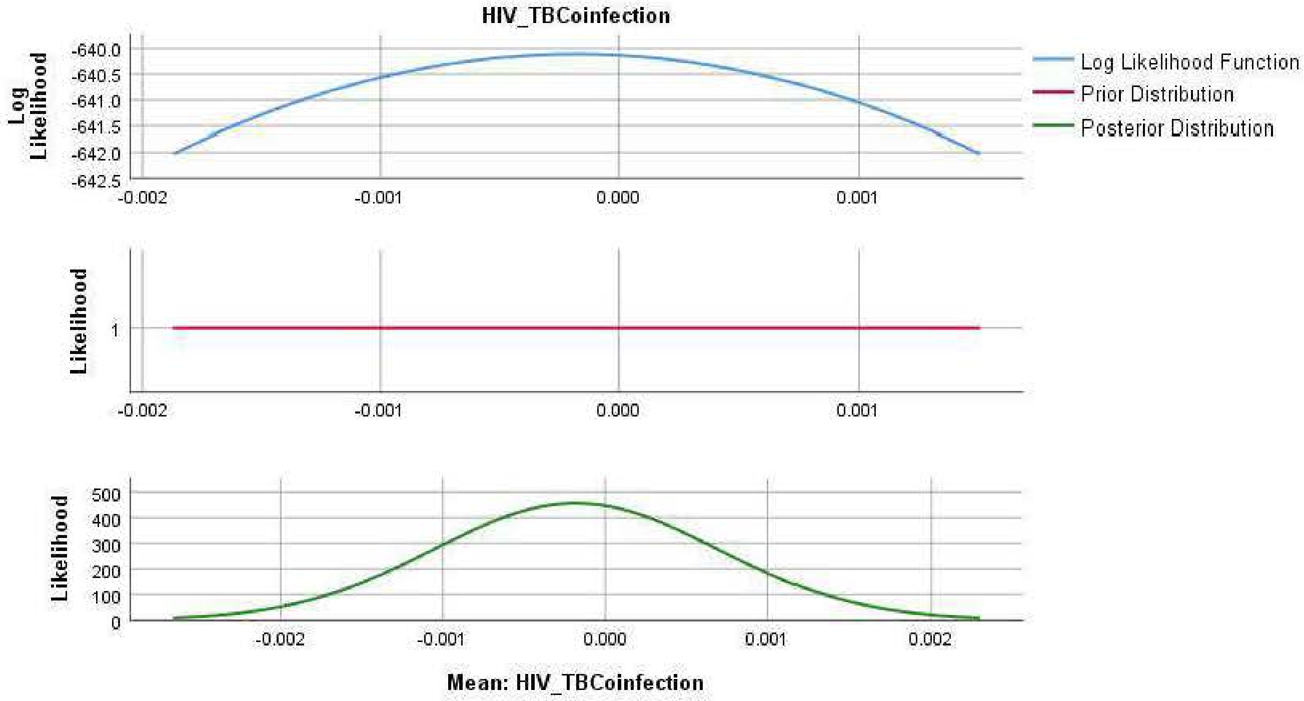


Figure 3: Graph showing likelihood functions for Number of Patients on Retention