DSPACE

https://dspace.org/

The spatial distribution of health facility viral suppression rates in the City of Matlosana Local Municipality, South Africa, 2011-2015

Ndlovu, Kudzayi

2024
https://hdl.handle.net/10539/37620

Downloaded from a DSpace repository. https://dspace.org/

The spatial distribution of health facility viral suppression rates in the City of Matlosana Local Municipality, South Africa, 2011-2015

By

Kudzayi Ndlovu

(Student Number: 1755809)

A research report submitted to the Wits School of Public Health, Faculty of Health Sciences, University of the Witwatersrand in partial fulfilment of the requirements for the degree of Master of Science Epidemiology and Biostatistics by combination of coursework and research

Supervisors

Professor. Kennedy Otwombe

Department of Epidemiology and Biostatistics,
School of Public Health, University of the Witwatersrand

Dr. Simbarashe Takuva

Division of Clinical Epidemiology, School of Health Systems and Public Health,
University of Pretoria

Johannesburg

June, 2023

DECLARATION

I declare that this research report is my own unaided work. It is being submitted for the degree of Master of Science Epidemiology and Biostatistics by combination of coursework and research at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Signature of candidate: 

Full name: Kudzayi Ndlovu

[14TH] day of [JUNE] in the year [2023].

DEDICATION

I dedicate this work to my family for the support that they have shown me during my time as a student, as I had to juggle between numerous personal commitments.

ACKNOWLEDGEMENTS

I would like to express my appreciation for the unwavering support from my supervisors Prof Kennedy Otwombe and Dr Simbarashe Takuva. Their input and mentorship meant that I learnt a lot during the course of my studies. The project would not have been possible without the support and assistance from the NHLS team whose data was used.

The team at the School of Public health and my colleagues played a pivotal role in various capacities which enabled me to conquer obstacles that I had to encounter during my time at the university. I will forever cherish the memories from the time spent at the University of Witwatersrand with its supportive staff.

LIST OF FIGURES

Figure 1: Map of the City of Matlosana Local Municipality showing heath facilities.....20

Figure 2: Mean Viral suppression rates by year 29

Figure 3: Facility viral suppression rates, 2011-2014 30

Figure 4: Facility viral suppression rates, 2015 and overall.....31

Figure 5: Ethics Clearance Certificate 44

Figure 6: Data Access Request.....45

LIST OF APPENDICES

Appendix 1.0: Study explanatory variables 40

Appendix 2.0: Overview of National Health Laboratory Service Corporate Data Warehouse Patient Linking.....41

Appendix 3.0: Project timelines.....43

Appendix 5.0: Data Access Request Letter 45

Appendix 6.0: Project management 46

Appendix 7.0: Dissemination plan 47

LIST OF TABLES

Table 2: Descriptive statistics for characteristics of participants in study facilities 26

Table 3: Viral suppression rates from 2011 to 2015 28

Table 4: Multivariate linear regression evaluating factors associated with viral suppression 32

Table 5: Study explanatory variables 40

hTable 6: Project timelines 43

Table 1: Budget for the study 46

ABSTRACT

Objective: The highly mobile population and a large immigrant population within City of Matlosana Local Municipality contributes to the high tuberculosis and HIV burden. This study seeks to use routine HIV laboratory data to compare facilities' viral suppression rates using geospatial techniques. Geospatial techniques give a visual presentation of the data including trends over time and patterns across facilities.

Design: This was a longitudinal study design based on secondary data obtained from the NHLS laboratory in Klerksdorp for 19 health facilities in City of Matlosana Local Municipality between 2011 and 2015.

Method: Program data extracted from the NHLS CDW database was used for the study. Data were collected each time a client comes for routine laboratory tests. The database contained data for all patients from the City of Matlosana Local Municipality who had their HIV care follow up tests done in the public sector between 2011 and 2015. Facility geocodes were used for mapping out suppression rates by facility and to present the data on a descriptive map. The choropleth map was computed using the functionality on QGIS that shows the dependent variable using color codes based on the corresponding percentage. This was done at baseline and over the follow up period to 2015 (spatiotemporal trends). STATA version 15 was used for the examination of the presence of clustering, where the Moran's Index was computed at 5% level of significance. The Simultaneous Autoregressive model was used to examine factors associated with suppression while adjusting for geocodes. Akaike Information Criterion was used for determining the best model fit.

Results: Of the total number of facilities assessed, 47% had the viral suppression rates between 60% and 70% while 21% of the facilities had their viral suppression ranging from 70 – 80%. The overall viral suppression rates were lowest in Klerksdorp hospital with 43.7% and the highest at Stilfontein municipal clinic with 78.2%. The viral suppression rates at Klerksdorp Hospital were the lowest throughout the entire study period. These rates started at 29.3% in 2011 and gradually rose over time, reaching a peak of 59.7% in 2015. The viral suppression rates across health facilities increased over time with Orkney and Stilfontein reporting highest performing. The yearly viral suppression rates showed a general increase in the viral suppression rates ranging from 69.4% in 2011 to a peak of 73.1% in 2012 and dipping in 2015 with a final rate of 72.1%. The local and global measure (Moran's I and Getis – Ordi) test for spatial autocorrelation was insignificant ($p=0.611>0.05$) and therefore spatial autoregressive models could not be used. However, results from general linear regression showed that viral suppression increased by 0.64 for every unit increase in age of clients tested (95%CI: 0.25 – 1.02; $p<0.01$). Furthermore, viral suppression decreased by 0.41 for every unit increase in white cell count (95%CI: -0.80 – -0.02; $p<0.05$).

Conclusion: Geospatial techniques can be employed to present a visual presentation of viral suppression rates, to identify trends over time and patterns across facilities. Results from the study showed that viral suppression rates ranged from 60% to 80% among the majority of facilities in the City of Matlosana Local Municipality. The lowest and highest viral suppression rates were at Klerksdorp hospital and Stilfontein municipal clinic respectively. Viral suppression rates across facilities were declining from 2012 to 2014, however there were sharp increases from 2011 to 2012 and from 2014 to 2015. White cell count and age for clients tested were significantly associated with viral suppression. In addition, there was no clustering of viral suppression rates in the City of Matlosana Local Municipality, implying that HIV viral loads varied across the health facilities.

Key words: Viral suppression rate, health facility, spatial distribution, CD4 count, GPS

LIST OF ACRONYMS

Acronym	Abbreviation
AIC	Akaike Information Criterion
ART	Antiretroviral Therapy
CD4	Cluster of Differentiation 4
GPS	Global Positioning System
HIV	Human Immunodeficiency Virus
KZN	KwaZulu Natal
MSM	Men Who Have Sex with Men
NHLS	National Health Laboratory Service
NGO	Non-Governmental Organization
NW	North West
PLHIV	People Living With HIV
QGIS	Quantum Geographical Information System
SAR	Simultaneous Autoregressive Model
SABSSMV1	The fifth South African National HIV prevalence, Incidence, Behavior and Communication Survey
UNAIDS	The Joint United Nations Program on HIV/AIDS
VS	Viral Suppression
WCC	White Cell Count

Table of Contents

DECLARATION ii
DEDICATION iii
ACKNOWLEDGEMENTS iv
LIST OF FIGURES v
LIST OF APPENDICES vi
LIST OF TABLES vii
ABSTRACT viii
LIST OF ACRONYMS x
CHAPTER ONE 13
1.0 Introduction 13
1.1 Background: 13
1.2 Literature review 14
1.3 Problem statement 17
1.4 Justification: 18
1.5 Research Question: 18
1.6 Hypothesis: 18
1.7 Main objective: 18
1.8 Specific objectives: 19
CHAPTER TWO 20
RESEARCH METHODOLOGY 20
2.1 Study design: 20
2.2 Study site and setting: 20
2.3 Source of Data 21
2.4 Sampling 22
2.5 Inclusion criteria 22
2.6 Power and sample size 22
2.7 Data Collection: 22
2.8 Description of Variables and Data Analysis 23
2.8.1 Definition of the outcome 23
2.8.2 Explanatory variables 23
2.8.3 Data management 23
2.8.4 Data analysis 24
2.9 Ethical considerations 25
CHAPTER THREE 26
RESULTS 26

3.1 Descriptive Statistics 26
3.2 Viral suppression rate by facility 27
3.3 Mean Viral suppression rates by year 28
3.4 Spatial distribution of facility VS 29
3.5 Test for Spatial Autocorrelation 32
3.6 Multivariate linear regression 32
CHAPTER FOUR 33
4.1 DISCUSSION 33
4.2 Limitations 35
5.0 REFERENCES 36
6.0 APPENDICES 40

CHAPTER ONE

1.0 Introduction

1.1 Background:

South Africa has the largest disease burden of human immunodeficiency virus(HIV) infection in the world with 7,7 million people infected(1). According to the Joint United Nations Program on HIV/AIDS(UNAIDS), the priority countries ought to attain the 90-90-90 targets by 2020(2). These state that 90% of all people living with HIV should be aware of their status and of those that are aware, 90% should be on treatment with the third 90 stating that 90% of those on treatment should be suppressed(2). The fifth South African National HIV prevalence, Incidence, Behavior and Communication Survey, 2017 (SABSSM V1) sought to give estimates of the key HIV indicators, with regards to the 90-90-90 targets. About 84.9% of all people living with HIV(PLHIV) between the ages of 15 and 64 were diagnosed (knew their status) with HIV, 70.6% of these were on antiretroviral therapy and of those who were on antiretroviral therapy (ART) 87.5% were suppressed. This attainment is an improvement towards the 90-90-90 targets from what was seen in 2012 where 32.9% of PLHIV were on ART with a 73.7% viral suppression rate(3).

The North West Province is one of the highest HIV burden provinces with a prevalence of 22.7% compared to the national prevalence of 20.6%. It has a population of about half a million people living with HIV. This accounts for about 7% of the country's population that is living with HIV(4). This variation in the provincial HIV burden has been noted to be due to disparity in the provincial budgetary allocations and this manifests in the differences in HIV testing rates, ART access and subsequently the viral suppression rates. The provinces that were noted to be spending more on HIV infected individuals were the Northern Cape and Free State and these had corresponding higher suppression rates(5). The low spenders were the Mpumalanga province and Gauteng with lower viral suppression rates (5).

This understanding is key in the evidence-based approach to public health decision making. Geospatial analysis can be used to inform and highlight priority public health facilities needing intervention in the North West Province in a bid to ensure that the 90-90-90 targets are met.

This will result in the scarce financial resources being put to efficient use through targeted interventions.

In 2015, the percentage of those on ART who were virally suppressed in South Africa was 78.4% (5). When considering suppression as a percentage of all PLHIV, the national suppression rate was noted to be 38.2% with a considerable provincial variation(3). This was far from the 73% target of overall suppression that is stipulated by the 90-90-90 targets. The percentage of PLHIV who are suppressed may be below the target of 73% but has been increasing owing to changes in the treatment guidelines. In 2012 it was 23.8% and steadily increased to 38.2% in 2015(3, 5). The change in the treatment guidelines over the years from CD4 based criteria to 'Test and Treat' meant that more people could be enrolled into care and treatment, increasing the second 90 and third 90 respectively to what is seen currently(1).

1.2 Literature review

1.2.1 Viral Suppression:

The use of ART seeks to suppress replication of the virus in those infected with HIV and consequently transmission to the uninfected commonly referred to treatment as prevention(6, 7). The attainment of the third ninety and realising the public health benefit is only possible through the monitoring of viral load in those on ART (2, 6). The monitoring of viral load has evolved over the years with lower limits of detection being realized due to improved assays that have been developed(6). This evolution has reflected in the way the South African treatment guidelines have changed from a suppression threshold of 400 to 50 copies per milliliter between 2015 and 2019 (8, 9). The varying thresholds may impact the suppression rates based on what cut off has been used. These varying thresholds make the comparison of some viral suppression studies difficult. An example is the population survey done in Kenya which applied a threshold of 550 copies/milliliter for suppression while the SABSSM V1 used a threshold of 1000 copies/milliliter as a definition of viral suppression(10).

1.2.2 Determinants of viral suppression

There are various factors that are associated with viral suppression. Viral suppression, as discussed earlier, is a function of effective antiretroviral therapy(8, 11). In those that are

already on ART, there are various factors that could potentially lead to a lack of suppression. One study, done in South Africa, looking at the impact of social capital on viral suppression and adherence showed that social capital is an important predictor of suppression in the community(12). This was noted to be related to the support structures for those who are HIV positive leading to increased self-confidence and consequently better adherence to ART (12). Several factors have been examined previously for association with viral suppression. In a study done in KwaZulu Natal(KZN) province, the following were found to predict poor viral suppression; gender, late initiation of ART as evidenced by advanced HIV infection stage, therapy related factors, probably due to adverse events with more failure being noted with stavudine and zidovudine as compared to tenofovir(13). The presence of adverse events impacts negatively on the ability of patients to adhere to the treatment and consequently encounter treatment interruptions.

In Ethiopia, determinants of viral suppression have been assessed and one of the predictors of poor suppression was a younger age group where the younger the person, the greater the likelihood of poor suppression(11). Like in the aforementioned study in KZN, defaulting of antiretroviral therapy was also noted even though the reasons why people default were not elicited; lack of condom use was noted to be another major predictor (10, 11, 14). There are occasions where therapy fails, despite good adherence, leading to an unsuppressed viral load. According to the South African national consolidated guidelines for treatment of HIV/AIDS, virological failure is diagnosed when the viral load is greater than 1000 copies/ml on at least two occasions more than two months apart despite good adherence(9).

1.2.3 Impact of location on suppression:

There are various aspects of medical phenomenon that have a geographical component to them(15). The viral suppression rates in South Africa are characterized by significant provincial variation where the Northern Cape Province has been noted to have the lowest at 52.8% while the Eastern Cape was noted to be the highest at 68%(16). According to a population based survey done in 2013, there was a suppression rate of 25.7% in the North West(NW) province using a threshold of 1000 copies/milliliter while a conservative cut off of 5000 copies/ml led to a 51.8% suppression rate(17). The NW suppression rates have, however, been generally improving over the years as evidenced by the current estimates of

57.8%(16). This has been motivated by the accelerated scale up of ART where the guidelines currently advocate for test and treat leading to increased ART coverage in the PLHIV and consequently more people suppressed from the population living with HIV(9).

There has been an assessment of community and individual level factors (including patient ZIP codes) in Philadelphia with regards to viral suppression and retention in care. The community level factors were found to be more significantly associated with clusters of poor viral suppression and poor retention in care than individual factors(18, 19). The association of poor suppression with community factors like poor economic status and gender disparities has been noted in studies done in various countries and the results have been consistent with an inclination towards poor suppression(11, 19).

The presence of a facility within a certain community does not always correlate with higher retention in care and viral suppression, people have been noted to travel long distances for care due to a variety of reasons which could be merely personal preference or influenced by facility factors and this, in most cases, is not a good predictor of suppression or retention (20). This means that there is always a discrepancy between a geographical catchment area and the actual catchment area. A study that used actual residential locations of HIV patients and the HIV care center location found an association between longer travel time to the facility and poor retention in care due to a lot of time spent travelling to the facility while suppression was not affected(14, 20). This could be attributed to the reason that when clients skip appointments, they will try and get replacement medication from a nearby facility leading to the findings of poor retention with good suppression. The use of patient residential locations as opposed to the ZIP codes comes with ethical issues since confidentiality is breached. Crime as a determinant of viral suppression was examined in one study in Chicago where the outcome found a significant association between high crime rate and poor viral suppression while retention in care was not associated with crime rate(14).

A provincial study done in the province of KwaZulu Natal noted clustering of predictors of poor suppression such as transactional sex, low levels of education, drug abuse, early sexual debut, multiple sexual partners and cohabiting(21). This is in line with Tobler's law that states that all objects are related in space, but nearer objects are more related to each other. In this

context, human behavior in certain communities will tend to be similar. For example, in low socio-economic status communities, there are bound to be more people engaging in transactional sex with high drug abuse rates among other behaviors.

1.3 Problem statement

HIV has transformed from being a generalized epidemic to being a localized epidemic within certain high risk populations which include certain communities; men who have sex with men(MSM), female sex workers, intravenous drug users and prisoners(2). The demographic profile of those at higher risk means that there is room for targeted interventions. The geospatial techniques have a role in the surveillance and monitoring of HIV programs and the detection of variations within provinces or districts(10). The South African HIV management guidelines recommend that viral load monitoring be employed as a marker of response to treatment in all PLWHA (9). There is limited use of routine laboratory data obtained from HIV care by decision makers, but there is room for improving usage of such data to direct policy and resource allocation right down to granular levels. To aid the accuracy of such data, there needs to be a strengthening of the ability of clinicians to order the viral load tests for patients at the right time, resulting in an efficient viral load cascade to ensure that progress towards the 90-90-90 targets is accurately monitored(22).

The progress towards attainment of the third ninety targets is assessed using a viral load threshold of 1000 copies/milliliter. The scale up of viral load monitoring has presented a unique opportunity to use this routine program data for assessing progress and evaluating the impact of this life saving therapy at a personal and population level. One way of presenting the data analyzed from the programs is through geographical information systems which make use of visualization to explain or highlight areas needing intervention which could be lagging behind with regards to attainment of 90-90-90 targets or merely lagging when compared to areas of similar socioeconomic characteristics (23). This has not been widely used in South Africa and Africa as whole.

1.4 Justification:

City of Matlosana Local Municipality has a highly mobile population and a large immigrant population due to the presence of mines. These contribute to the high tuberculosis and HIV burden seen in the district and this has seen the district being identified as a priority district for HIV and TB programs (24). This study seeks to use routine HIV laboratory data at National Health Laboratory Service (NHLS) level to compare facilities' viral suppression rates using geospatial techniques. The viral load threshold to be used is 1000 copies/ml because the study uses program data. This is a practice that is not used widely in relation to routine laboratory data. This would be an opportunity to showcase a wider usage of the data and further examine any potential associations using laboratory markers and health facility geospatial points as opposed to patient geospatial points for ethical reasons. The application of the geospatial techniques gives a visual presentation of the data (hot and cold spots) including trends over time and patterns across facilities in the district. This will be used by policy makers and program managers to inform targeted actions taken at facility level to address the problem(s) of poor suppression rates as they align to the 90-90-90 targets(2). This approach has not been fully utilized in the low to middle income countries and this presents opportunities on the usage of geographic information systems in HIV programming(23).

1.5 Research Question:

What is the spatial distribution of health facility viral suppression rates in the City of Matlosana Local Municipality, North-West province, South Africa during the period 2011 - 2015?

1.6 Hypothesis:

Null Hypothesis: The viral suppression rates are similar in space across different health facilities in the City of Matlosana Local Municipality.

1.7 Main objective:

To use geospatial methods to highlight differences in facility rates of viral suppression and associated factors based on routine laboratory data.

1.8 Specific objectives:

1.8.1. To use geospatial techniques to describe viral suppression rates across different health facilities in City of Matlosana Local Municipality between 2011 and 2015

1.8.2. To determine the presence of geospatial clustering of viral suppression rates in City of Matlosana Local Municipality between 2011 and 2015.

1.8.3. To determine factors associated with viral suppression, adjusting for geocodes at baseline, final visit and overtime(spatiotemporal) in City of Matlosana Local Municipality

CHAPTER TWO

RESEARCH METHODOLOGY

2.1 Study design:

A secondary analysis of longitudinal data was undertaken using routine HIV care data (secondary data) obtained from the NHLS laboratory in Klerksdorp which serviced health facilities in City of Matlosana Local Municipality between 2011 and 2015. Data were collected each time a client comes for routine laboratory tests. The study made use of shapefiles covering the district while the clinic positions were obtained from google earth.

2.2 Study site and setting:

The City of Matlosana Local Municipality is a Category B municipality situated within the Dr Kenneth Kaunda District in the North West Province (Fig 1). The City of Matlosana has a total population of 438 486 people, of whom 103 407 (92%) are urbanised and 35 079 (8%) are rural. The study population comprised of participants from the City of Matlosana Local Municipality municipality whose viral load tests were done at the NHLS in Klerksdorp. These were from the 19 health facilities within the City of Matlosana Local Municipality, but may have people who seek care there without necessarily staying there.

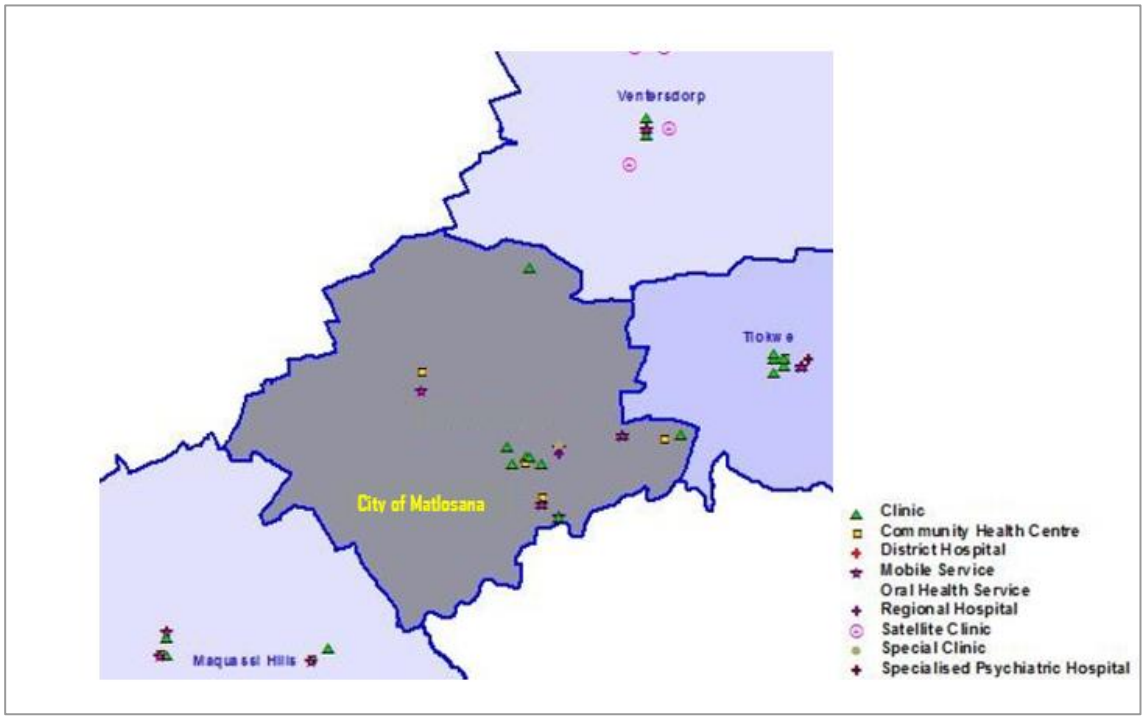

Figure 1: Map of the City of Matlosana Local Municipality showing heath facilities

The population consisted of clients on ART regardless of whether it was first or second line as the laboratory data did not capture this information. When patients test positive for HIV, they are initiated on first line therapy having followed the latest treatment algorithms(8, 9). The criteria for initiation on ART have changed over time, it was initially based on CD4 count of 200 cells/milliliter and then changed to 350 cells/milliliter and subsequently 500 cells/milliliter. The same CD4 count measurement was used for monitoring response to treatment in stable patients unless a patient was improving clinically. This has now been replaced by an objective way of measuring response to therapy which is viral load measurement(8). The monitoring of viral load was according to the South African HIV treatment guidelines where those with viral loads above 1000copies/milliliter eight weeks apart, despite good adherence, were switched to second line(8). Stable or suppressed patients have their viral load measurements done once a year, while those who were unsuppressed have it done at least twice a year according to the treatment guidelines(6, 8). Kenneth Kaunda district is a mining area characterized by a high HIV and TB burden mentioned previously. The health facilities in question were the public health facilities which range from primary to tertiary health facilities.

2.3 Source of Data

The National Health Laboratory Service Corporate Data Warehouse (CDW) is the largest diagnostic pathology laboratory service in South Africa, with 349 laboratories across the nine provinces supporting the 4000 clinics and 400 hospitals in the South African public healthcare system(25). The CDW is a repository for all public sector laboratory measurements in South Africa (over 80% of all facilities, >43 million people), which includes all HIV-related laboratory tests conducted across public health laboratories since 2004. All HIV-related laboratory tests, that is CD4+ cell counts, viral load measurements conducted from 2004 were captured in the system, and this data from the City of Matlosana Local Municipality, North-West province was extracted for analysis.

The extracted dataset was created from a number of steps. A laboratory request form with demographic data accompanies each specimen and was recorded onto the laboratory information system. Once testing was complete, all available results were recorded alongside the demographic data. As laboratory tests can and do relate to the same individual having

multiple tests over time and between sites, data linkage was performed using a probabilistic algorithm to determine the records that relate to the same individual. In probabilistic record linkage, personal identifiers are used together to determine how likely every single pair of records relates to the same individual(26-28). First, records were directly linked in which surnames, first names and dates of birth were an exact match. After this, probabilistic matching, which uses an algorithm to assess the probability that a set of records were linked to the same individual, was conducted. The resulting probabilities were weighted and averaged, according to the frequency within the population, to provide a final score on which matching was based and sequentially numbered(29). A database was constructed to enable a single-record-per-person within a year and a chronological laboratory history for each person across years. An overview of this patient linking algorithm is supplied as Appendix 3.

2.4 Sampling

Klerksdorp serviced 19 health facilities in City of Matlosana Local Municipality. Total sampling was used from 19 facilities from the City of Matlosana Local Municipality with patients in the database who had their viral load tests done.

2.5 Inclusion criteria

Participants who were 18 years of age and above with at least one viral load result between 2011 and 2015 were included in the study. They were included regardless of treatment status.

2.6 Power and sample size

Power for this study was determined using the one-sample proportion test for cluster-randomized design available in STATA version 15 (STATA Corp. College Station, TX, USA). The study enrolled at least 78,000 HIV infected people from 19 facilities to detect a minimum difference of 6.5%. A power assumption of 80% was made together with an intracluster correlation coefficient of 0.2. The level of significance was 5% while a null proportion of 87.5% was used.

2.7 Data Collection:

Program data extracted from the NHLS CDW database was used for the study. The database contained data for all patients from City of Matlosana Local Municipality who had their

various HIV care follow up tests done in the public sector between 2011 and 2015. The data included the basic demographic information and the actual laboratory parameters that were measured routinely as part of standard HIV care.

2.8 Description of Variables and Data Analysis

2.8.1 Definition of the outcome

The primary outcome of this study was facility viral suppression rate and this was a continuous variable. The viral suppression rates incorporated the facility individual viral loads which were below the threshold at facility level. The threshold that was used in this study was a viral load of 1000 copes/milliliter. This means that an individual was defined as suppressed if the viral load was below 1000 copies/milliliter and unsuppressed if viral load was greater than 1000 copies/milliliter. The cut off was chosen based on the South African HIV treatment guidelines and on the basis of the 90-90-90 targets(2, 9). These were aggregated to give the suppression rates by facility.

2.8.2 Explanatory variables

The list of variables used in the study are summarized in appendix 2.

2.8.3 Data management

Data was deidentified and use of identifiers was done as discussed in the preceding section to protect the privacy of the patients that were in the database. The data was password protected for security purposes with only the investigator and the supervisors having access to the password.

This was routine data and data cleaning was employed to ensure that there were no duplicates in the dataset, outliers or any other data related problems. The participants with missing values for the variables under study were excluded from the analysis. The data will be kept for two years after the study has been published.

2.8.4 Data analysis

STATA version 15 (STATA Corp., College Station, TX, USA) was used for all statistical analysis. Statistical significance was based on 95% confidence interval (5% level of significance) to assess the strength of association between different variables. Data transformation was employed in the analysis when necessary.

The descriptive analysis was done by facility with detail percentage composition of the explanatory variables. The coordinates for the various facilities were obtained using google earth and the choropleth map done as outlined under objective 1.

Analysis by objective was as follows:

Objective 1 To use geospatial maps to describe viral suppression rates across different health facilities in City of Matlosana Local Municipality

Analysis

The facility viral suppression rate was the outcome variable using the threshold of 1000 copies/ml. The proportion of suppressed individuals by facility contributed towards that facility's suppression rate. The values for proportion of suppressed individuals ranged between 0 and 100%. The facility geocodes were used for mapping out suppression rates by facility and to present the data on a descriptive map. QGIS was used to merge the excel (CSV) format with facility coordinates, suppression rates and the shapefile with wards in the Dr Kenneth Kaunda municipality. The choropleth map was done using the functionality on QGIS that shows the dependent variable using color codes based on the corresponding percentage. This was done at baseline and over the follow up period to 2015 (spatiotemporal trends).

Objective 2 To determine the presence of geospatial clustering on facility viral suppression rates.

Analysis

The examination of the presence of clustering was done using STATA version 15 where the Moran's Index was computed at 5% level of significance.

Objective 3 To determine factors associated with viral suppression adjusting for geocodes at baseline, final visit and overtime (spatiotemporal)

Analysis

The test for spatial autocorrelation was done prior to running our model using the "estat-moran" commands on STATA after running ordinary regression on STATA. The Simultaneous Autoregressive (SAR) model was used to examine factors associated with suppression while adjusting for geocodes.

The SAR model used was a spatially autoregressive error model where;

$Y=X\beta + \lambda W\mu+\varepsilon$

The response variable is denoted by Y which is the viral suppression while X denotes a vector that is linearly dependent on the study of independent variables (hemoglobin, white cell count and CD4 count). W is the spatial weights matrix while λ is the scalar that measures the spatial correlation in the errors. μ gives us the spatial correlated errors and Ɛ gives the error term.

Ordinary linear regression analysis was applied to determine factors associated with viral suppression. Selected factors were identified and quantified through stepwise linear regression analysis (cut-off, $p<0.10$). Significant factors were re-analysed in a new model Akaike Information Criterion (AIC) was used for determining the best model fit. Akaike's information criterion (AIC) is a tool for assessing model fit. A model with a smaller AIC value indicates a "better" model.

Cook's Weisberg test for constant variance and Ramsey RESET test for linearity were employed for linear regression diagnostics. In addition, variance inflation factor (VIF) and the Shapiro-Wilk test were applied to test for collinearity and normality.

2.9 Ethical considerations

Ethical clearance was received from the University of Witwatersrand human research ethics committee (See appendix 4.0, M210304), the NHLS laboratory in Johannesburg and the Center for HIV and STIs at NICD-NHLS (See appendix 4.0, M210304). The patient names or identifiers were removed to ensure confidentiality of the participants and there was no use of the patient residential geolocation.

CHAPTER THREE

RESULTS

This section provides results with the aims of addressing the objectives of the study. The chapter begins by outlining the overall descriptive statistics. The subsequent section discusses the variation of viral suppression rates over time of the respective health facilities. The penultimate section then incorporates the spatial techniques to outline the variation in space. The chapter ends with tests for spatial autocorrelation and multivariate linear regression.

3.1 Descriptive Statistics

Table 2 shows the characteristics of the 19 facilities in the study. There were 17 clinics and two hospitals from the study area. The age of participants was lowest at Klerksdorp Hospital 34(13) and Klerksdorp Prison 34(11). On the other hand, the highest average age of participants, was highest at Stilfontein clinic 42(15).

The Orkney Municipal Clinic and Khuma Clinic had the lowest white cell count (WCC) of 5.2(2.6) and 5.2(2.8) respectively. On the other hand, Klerksdorp Hospital had the highest WCC of 6.8(4.1), making it the only facility with an average WCC above 6. The average haemoglobin levels were similar across most facilities, except for Klerksdorp Hospital, which had the lowest haemoglobin levels of 10.9(3.3)g/dl. Conversely, Klerksdorp Prison had the highest average haemoglobin levels of 14.6(2.1)g/dl. It was the only facility with an average haemoglobin above 13, while 12 facilities had average haemoglobin levels between 12 g/dl and 13 g/dl.

The average platelet count, Klerksdorp Prison had the lowest count of 221(86), while Park Street Clinic had the highest count of 270(111). Ten facilities had average platelet counts between 250 and 260, while six facilities had counts between 260 and 270. Tshepong Hospital had the lowest average CD4 count of 358(391), whereas Alabama had the highest average CD4 count of 469(357). Fourteen facilities had average CD4 counts above 400, while the remaining facilities had counts between 300 and 400.

Table 1: Descriptive statistics for characteristics of participants in study facilities

Facility_Name	Age_Tested_Years*	WCC*	Haemoglobin*	Platelets*	Cd4_Count*
Alabama Clinic	39(14)	5.9(2.8)	13(2.9)	256(106)	469.0(357)
Botshabelo Clinic	41(16)	5.48(2.8)	12.6(2.6)	250(107)	415(335.4)
Delekile Khoza Clinic	35(13)	5.9(3.4)	12.5(3.7)	257(130)	382(365)
Empilisweni Clinic [Jouberton]	38(15)	5.6(2.5)	12.6(2.4)	264(105)	457(339.9)
Grace Makhomo Clinic	40(16)	5.5(2.7)	12.9(2.5)	268.5(109)	438(370.4)
Jouberton Clinic	39(16)	5.8(2.8)	13.2(2.5)	265(104)	434(346.2)
Kanana Clinic [Orkney]	40(15)	5.6(2.83)	13.1(2.5)	261.5(106)	433(342.9)
Khuma Clinic	39(15)	5.2(2.8)	12.5(2.7)	256.5(114.5)	395(325)
Klerksdorp Hospital	34(13)	6.8(4.1)	10.9(3.3)	257(122)	364(347.5)
Klerksdorp Prison	34(11)	5.4(2.8)	14.6(2.1)	221(86)	468(349)
Majara Sephapha Dist Clinic	40(17)	5.8(3.4)	12.4(2.6)	258(93)	414(323)
Marcus Zenzile Clinic	38(16)	5.3(2.7)	12.2(2.6)	259(111)	429(324)
Nm Pretorius Gateway Clinic	41(15)	5.9(3.0)	12.4(2.8)	264(90)	458 (382.3)
Orkney Municipal Clinic	38(14)	5.2(2.6)	13(2.2)	250(105)	410.5(301.5)
Park Street Clinic	41(14)	5.8(3.1)	12.6(2.8)	270(111)	457(349.0)
Stilfontein Municipal Clinic	42(15)	5.4(2.9)	12.8(2.3)	251(111)	458.5(343)
Tigane Clinic	38(15)	5.7(2.9)	13(2.9)	247(100)	395(354.4)
Tshepong Hospital	41(14)	5.9(3.5)	12.5(3.4)	264(129.0)	358(391)
Tsholofelo Clinic	41.0(15.0)	5.96 (2.9)	12.85(2.9)	254(118)	442 (318.0)

**The figures are presented as Median(IQR)*

3.2 Viral suppression rate by facility

The viral suppression rates (VS) were computed by facility with the lowest overall VS rate being from Klerksdorp hospital at 43.67% and the highest at Stilfontein municipal clinic at 78.21%. Klerksdorp hospital had the lowest VS rates across the entire study period starting from 29.33%, gradually increasing over the study period to a peak of 59.67% in 2015 (Table 3).

There were 14 facilities that had an overall VS rate between 70% and 80%. Klerksdorp hospital had the lowest rate of 43.67% (Table 3). In the year 2011, three health facilities (Klerksdorp Hospital, Klerksdorp Prison and Tshepong Hospital) had VS rates below 50% with the number dropping to one throughout the study period. This number coincided with the general improvement in the VS rates.

The VS rates across the health facilities generally increased over time with the best performing facilities noted to be Orkney and Stilfontein. The yearly VS rates showed a general

increase starting from 69.3% in 2011 to a peak of 73.1% in 2012 and a decrease to 70.3% and 69.1% in 2013 and 2014 respectively, with a final rate 72.1% in 2015 (Table 3).

Table 2: Viral suppression rates from 2011 to 2015

Facility_Name	2011*	2012*	2013*	2014*	2015*	Overall*	X Coordinates	Y Coordinates
Alabama Clinic	72.98	76.27	72.24	74.33	73.29	73.83	26.59043	-26.87193
Botshabelo Clinic	68.90	76.49	74.78	71.08	70.56	72.05	26.84639	-26.85447
Delekile Khoza Clinic	74.07	71.88	65.12	62.26	64.20	64.83	26.4163921	-26.7381858
Empilisweni Clinic [Jouberton]	66.93	73.64	68.85	69.87	71.64	70.41	26.603554	-26.887151
Grace Makhomo Clinic	71.66	76.17	69.99	69.60	70.75	71.15	26.6352351	-26.9560641
Jouberton Clinic	69.77	75.26	76.93	68.91	69.74	71.59	26.60673	-26.89532
Kanana Clinic [Orkney]	67.57	77.26	72.33	74.27	77.06	74.22	26.6755499	-26.98778
Khuma Clinic	53.76	75.00	71.03	72.92	71.74	70.39	26.77452	-26.84282
Klerksdorp Hospital	29.33	33.29	39.97	45.18	59.67	43.67	26.66388	-26.87788
Klerksdorp Prison	39.04	75.00	60.63	50.00	70.12	56.18	26.6620286	-26.9090983
Majara Sephapha Dist Clinic	67.43	78.56	72.41	64.63	74.14	71.78	26.61859	-26.96753
Marcus Zenzile Clinic	61.14	75.62	75.78	74.60	76.52	74.78	26.8192715	-26.8542603
Nm Pretorius Gateway Clinic	74.88	83.21	77.76	70.08	77.09	76.19	26.614876	-26.885815
Orkney Municipal Clinic	59.74	76.10	83.49	82.66	76.82	77.80	26.6755499	-26.98778
Park Street Clinic	67.35	79.55	78.62	76.23	78.46	76.97	26.66792	-26.83484
Stilfontein Municipal Clinic	64.91	77.05	81.59	79.82	84.23	78.21	26.77502	-26.84464
Tigane Clinic	56.67	68.34	63.49	64.59	65.69	64.65	26.42192	-26.73537
Tshepong Hospital	48.94	60.87	56.52	65.17	63.85	59.34	26.61646	-26.885949
Tsholofelo Clinic	69.93	79.67	73.83	76.03	74.93	75.13	26.588967	-26.8918562
Yearly average across facilities	62.37	73.12	70.28	69.06	72.13	69.64		

**The figures are presented as averages*

3.3 Mean Viral suppression rates by year

Results showed a sharp increase in the overall mean VS across all facilities from 2011 to 2012 (Figure 2). However, there was a decline from 2012 to 2014 and a sudden increase as from 2014.

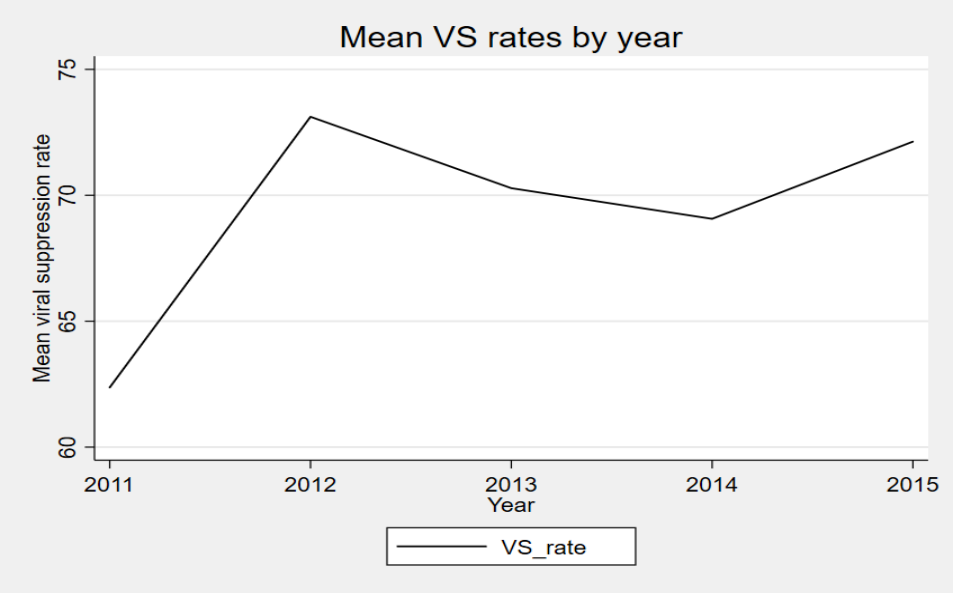

Figure 2: Mean Viral suppression rates by year

3.4 Spatial distribution of facility VS

Figures 2-3 shows the spatial distribution of the individual facility VS rates. Results showed that distribution of facilities was mostly around a few enumeration areas, this was as a result of the urban areas that were found in those particular areas. Of the total number of facilities, 47% of the facilities have the VS rates between 60% and 70% while 21% of the facilities had VS rates that ranged from 70% to 80%. There was one health facility that had a VS rate between 40% and 50%. In the map it is in very close proximity to a site that has between 70% and 80% suppression.

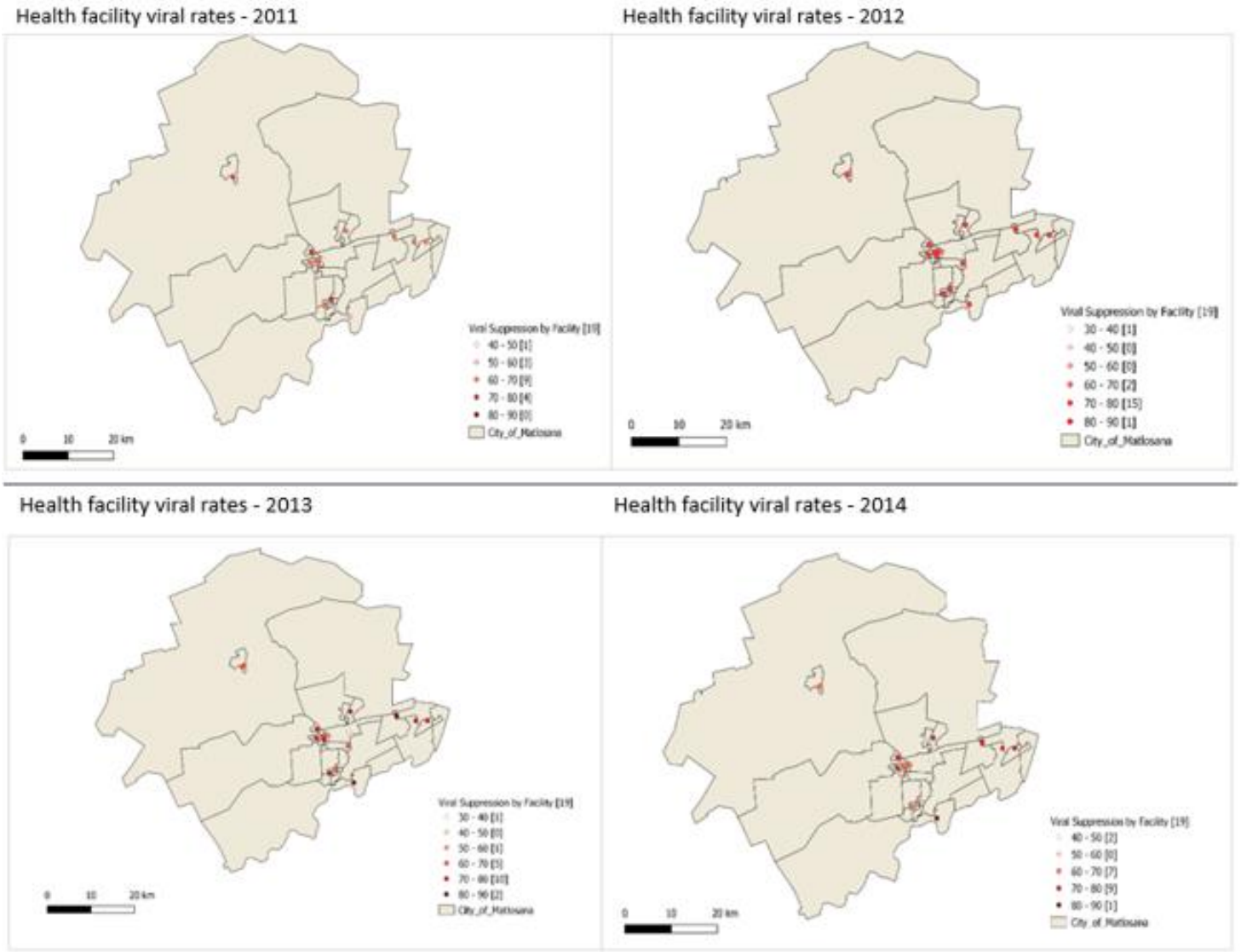

Figure 3: Facility viral suppression rates, 2011-2014

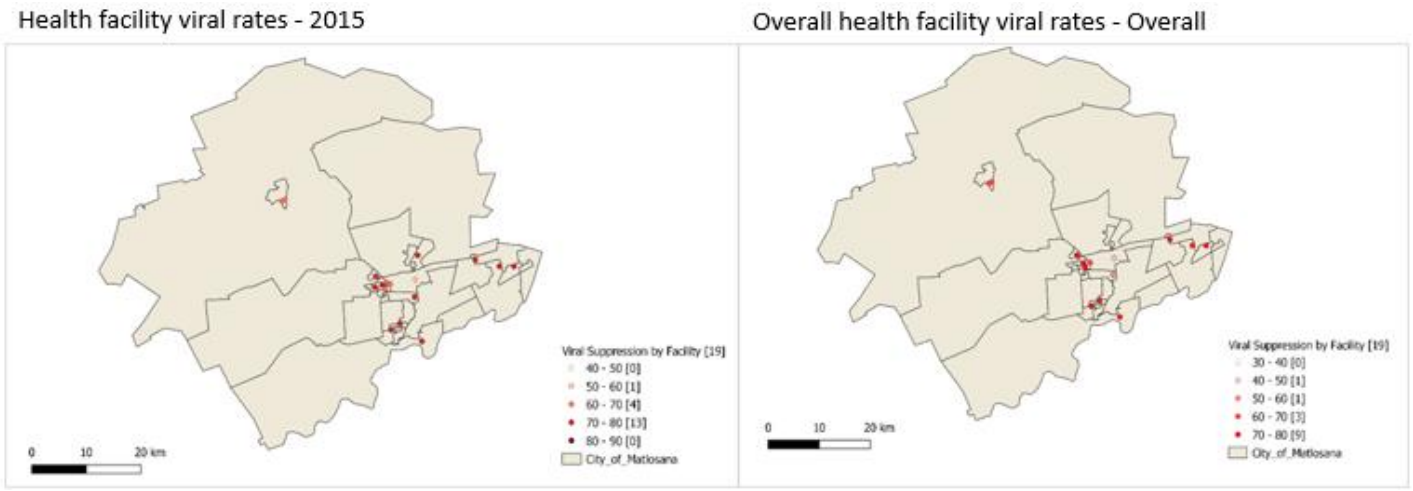

Figure 4: Facility viral suppression rates, 2015 and overall

3.5 Test for Spatial Autocorrelation

The local and global measure (Moran's I and Getis – Ordi) test for spatial autocorrelation was insignificant (p=0.611>0.05) and therefore spatial autoregressive models could not be used. However multivariate linear regression analysis was further employed to determine factors associated with viral suppression. Cook's Weisberg and the Ramsey RESET test showed evidence of constant variance and linearity respectively (Table 4). In addition, the Shapiro-Wilk test showed evidence of normal distribution for the predicted residual values.

3.6 Multivariate linear regression

Both forward and backward stepwise linear regression at cut-off 0.10 retained 'age tested' and 'white cell count (WCC)'. The Akaike's information criterion (AIC) test for best model fit was in favour of the final model with retained factors after stepwise linear regression. Results from the final model showed that viral suppression increased by 0.64 for every unit increase in age of clients tested (95%CI: 0.25 – 1.02; p<0.01). Viral loads decreased by 0.41 for every unit increase in white cell count (95%CI: -0.80 – -0.02; p<0.05), (Table 4).

Table 3: Multivariate linear regression evaluating factors associated with viral suppression

Site Characteristic	**Coef**	**Std.Err**	**p-value**	**95%CI**
Age tested	0.638	0.181	*	0.253; 1.023
White Cell Count	-0.410	0.185	**	-0.801; -0.019
Akaike's information criterion (AIC)	Null Model: 115.20		Final Model: 111.91	
Cook's Weisberg test	p = 0.842			
Shapiro-Wilk test	p = 0.742			
Ramsey RESET test	p = 0.956			

Key

*** Significant, p<0.001

** Significant, p<0.01

* Significant, p<0.05

CHAPTER FOUR

4.1 DISCUSSION

The objective of the study was to investigate the spatial distribution of viral suppression rates in health facilities within the City of Matlosana Local Municipality from 2011 to 2015. The study utilized a suppression threshold of 1000 copies/ml. Results indicated that approximately 15% of the total antiretroviral therapy (ART) provision during the study period was attributed to private sector and non-governmental organization (NGO) programs. The exclusion of data from non-public sector sources may introduce the possibility of either underestimating or overestimating the viral suppression rates. Furthermore, since the study primarily relied on laboratory data, there was a lack of information concerning individuals' treatment regimens (first- or second-line) and the presence of any co-occurring opportunistic illnesses.

The 90-90-90 targets aim to achieve 90% of all people living with HIV (PLHIV) to know their HIV status, 90% of all people with diagnosed HIV infection to receive antiretroviral therapy (ART), and 90% of all people receiving ART to be virally suppressed (90-90-90 targets) (1). These 90-90-90 targets were meant to be attained by 2020 and the results showed an increase in the viral suppression rates over time from 2011 to 2015 (Figure 2)

The increase in suppression rates across health facilities over time is in line with the routinization of viral load monitoring. The viral load coverage rates for countries in the sub-Saharan Africa in 2013 were noted to be generally less than 10% with the exception of a few countries whose coverage was significantly higher(2). The scale up of viral load testing in most of the sub-Saharan countries aided by PEPFAR improved the coverage rates to at least 50%, this improvement also coincided with a significant improvement in the viral load suppression rates. This means that the viral load test would be applied to all patients regardless of whether they are not adherent to antiretroviral therapy.

Furthermore, the positive trend in viral load performance and coverage over time can be attributed to accessibility to testing and treatment, no severe opportunistic infection and no treatment interruptions, good nutrition status, and peer support (30). However, the ADreSS

study reported a viral load coverage of approximately 38.7%, falling significantly short of the targeted goal of 90%(3). This low coverage of viral load means that there are a lot of patients with potentially high viral loads who are potentially transmitting the virus undermining the HIV prevention efforts and further contributing to the morbidity and mortality from HIV. This highlights the necessity for additional interventions and strategies to enhance viral load testing and achieve optimal coverage. Successful viral suppression is strongly influenced by medication adherence and clinic support (31). In rural communities, healthcare access, stigma, and patient engagement have been identified as potential barriers to viral load monitoring and suppression (32). Addressing these challenges is crucial to improve outcomes in these settings.

The spatial analysis conducted in this study did not identify any significant clustering of viral suppression rates within the City of Matlosana Local Municipality. The majority of health facilities were located in urban areas, while some were scattered across mining and farming regions. It is important to consider that the small sample size of facilities may have influenced these findings (33). Examining the viral suppression rates across the health facilities, shows that many of them fell within a similar range. This consistency can be attributed to the shared knowledge and clinical approach to HIV care, as these facilities receive support from the same district office or partner organizations (34). Furthermore, a notable finding in this study was the persistently low viral suppression rates at Klerksdorp Hospital, consistently below 60%. When compared to other sites over time, it had the lowest rates of viral suppression. Psychiatric illness is among factors associated with the presence of persistent viremia in HIV-infected patients (35).

During the study period, the City of Matlosana struggled to meet the 90-90-90 targets, with certain facilities exhibiting suppression rates below 50%. Notably, Klerksdorp Prison and Tshepong Hospital recorded the lowest rates of viral suppression. The low rates of viral suppression observed could be attributed to the highly mobile nature of the population in the area. The final analysis demonstrated a statistically significant association between age and viral suppression, indicating that for each unit increase in age of the clients tested, viral suppression increased by 0.64. Findings from a previous study also revealed correlation between age and viral suppression, the results indicated a positive association with increasing

age. (36). Furthermore, our study revealed an inverse relationship between white cell count and viral loads, with viral loads decreasing by 0.41 for every unit increase in white cell count. This outcome aligns with a previous study that similarly observed a reduction in viral loads with increasing white cell count (37).

4.2 Limitations

The study was based on secondary data from routine HIV care program, this resulted in a smaller number of facilities and a smaller study area. The study area was meant to be Dr Kenneth Kaunda district municipality and eventually it was just the City of Matlosana which is one of the three municipalities that make up the district. Since the study design was longitudinal and based on secondary data, laboratory values for certain patients were missing due to loss of follow up, transfer to another facility and even death.

The data was from 2011 to 2015, the viral load measurement or request by the practitioners has improved over time and this may make comparisons of the various years difficult due to missing viral load values in the earlier parts of the study. There may be missing values for certain variables like age, clinic or the various laboratory measures and that may result in measurement error of the association. Similarly, the study's limitation lies in its inability to explore community and individual factors, such as crime rate, distance to healthcare facilities, socioeconomic status, and education, which have been highlighted in the literature as important contributors to the lack of viral suppression (38). This limitation arises due to the available data being confined to laboratory markers without comprehensive information on these factors.

Furthermore, it is important to note that the participants seeking care at specific facilities may not necessarily reside within the catchment areas of those facilities, as ZIP codes were not utilized for geolocation purposes. Furthermore, NHLS data are specimen-based, it does not contain ART-specific data, presenting challenges in determining whether patients were receiving ART or not. In the context of HIV treatment and monitoring guidelines during the study period, the presence of a viral load measurement as a proxy was used for ART status after deduplication.

5.0 REFERENCES

1. (2018) HSRCH. The Fifth South African National HIV Prevalence, Incidence, Behaviour and Communication Survey. Cape Town: HSRC; 2018.
2. UNAIDS. 90-90-90 An ambititious treatment target to help end the AIDS epidemic. Geneva: UNAIDS; 2014.
3. Takuva S, Brown AE, Pillay Y, Delpech V, Puren AJ. The continuum of HIV care in South Africa: implications for achieving the second and third UNAIDS 90-90-90 targets. AIDS. 2017;31(4):545-52.
4. Lippman SA, El Ayadi AM, Grignon JS, Puren A, Liegler T, Venter WDF, et al. Improvements in the South African HIV care cascade: findings on 90-90-90 targets from successive population-representative surveys in North West Province. J Int AIDS Soc. 2019;22(6):e25295-e.
5. Johnson LF, Dorrington RE, Moolla H. Progress towards the 2020 targets for HIV diagnosis and antiretroviral treatment in South Africa. 2017. 2017;18(1).
6. Drain PK, Dorward J, Bender A, Lillis L, Marinucci F, Sacks J, et al. Point-of-Care HIV Viral Load Testing: an Essential Tool for a Sustainable Global HIV/AIDS Response. Clinical Microbiology Reviews. 2019;32(3):e00097-18.
7. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med. 2011;365(6):493-505.
8. Health NDo. 2019 ART Clinical Guidelines for the Management of HIV in Adults, Pregnancy, Adolescents, Children, Infants and Neonates. In: DOH, editor. South Africa2019. p. 26.
9. NDOH. National Consolidated Guidelines In: NDOH, editor. 2015. p. 75.
10. Cherutich P, Kim AA, Kellogg TA, Sherr K, Waruru A, De Cock KM, et al. Detectable HIV Viral Load in Kenya: Data from a Population-Based Survey. PLoS One. 2016;11(5):e0154318-e.
11. Bayu B, Tariku A, Bulti AB, Habitu YA, Derso T, Teshome DF. Determinants of virological failure among patients on highly active antiretroviral therapy in University of Gondar Referral Hospital, Northwest Ethiopia: a case-control study. HIV/AIDS (Auckland, NZ). 2017;9:153-9.

12. Mukoswa GM, Charalambous S, Nelson G. The association between social capital and HIV treatment outcomes in South Africa. PLoS One. 2017;12(11):e0184140.
13. Hunt GM, Dokubo EK, Takuva S, de Oliveira T, Ledwaba J, Dube N, et al. Rates of virological suppression and drug resistance in adult HIV-1-positive patients attending primary healthcare facilities in KwaZulu-Natal, South Africa. The Journal of antimicrobial chemotherapy. 2017;72(11):3141-8.
14. Ridgway JP, Almirol EA, Schmitt J, Schuble T, Schneider JA. Travel Time to Clinic but not Neighborhood Crime Rate is Associated with Retention in Care Among HIV-Positive Patients. AIDS and Behavior. 2018;22(9):3003-8.
15. Kandwal R, Garg PK, Garg RD. Health GIS and HIV/AIDS studies: Perspective and retrospective. Journal of Biomedical Informatics. 2009;42(4):748-55.
16. HSRC. THE FIFTH SOUTH AFRICAN NATIONAL HIV PREVALENCE, INCIDENCE, BEHAVIOUR AND COMMUNICATION SURVEY, 2017 (SABSSM V1). HSRC; 2018.
17. Lippman SA, Shade SB, El Ayadi AM, Gilvydis JM, Grignon JS, Liegler T, et al. Attrition and Opportunities Along the HIV Care Continuum: Findings From a Population-Based Sample, North West Province, South Africa. J Acquir Immune Defic Syndr. 2016;73(1):91-9.
18. Eberhart MG, Yehia BR, Hillier A, Voytek CD, Blank MB, Frank I, et al. Behind the cascade: analyzing spatial patterns along the HIV care continuum. J Acquir Immune Defic Syndr. 2013;64 Suppl 1(0 1):S42-S51.
19. Eberhart MG, Yehia BR, Hillier A, Voytek CD, Fiore DJ, Blank M, et al. Individual and community factors associated with geographic clusters of poor HIV care retention and poor viral suppression. J Acquir Immune Defic Syndr. 2015;69 Suppl 1(0 1):S37-S43.
20. Terzian AS, Younes N, Greenberg AE, Opoku J, Hubbard J, Happ LP, et al. Identifying Spatial Variation Along the HIV Care Continuum: The Role of Distance to Care on Retention and Viral Suppression. AIDS and Behavior. 2018;22(9):3009-23.
21. Wand H, Ramjee G. Spatial clustering of “measured” and “unmeasured” risk factors for HIV infections in hyper-endemic communities in KwaZulu-Natal, South Africa: results from geoadditive models. AIDS Care. 2015;27(11):1375-81.
22. Euvrard J, Schulz T, Hilderbrand K, Bosland M, Osler M, Boulle A, et al. How accurately do routinely reported HIV viral load suppression proportions reflect progress towards the 90-90-90 target in the population on antiretroviral treatment in Khayelitsha, South

Africa? South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde. 2019;109(3):174-7.

23. Lilian RR, Grobbelaar CJ, Hurter T, McIntyre JA, Struthers HE, Peters RPH. Application opportunities of geographic information systems analysis to support achievement of the UNAIDS 90-90-90 targets in South Africa. South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde. 2017;107(12):1065-71.
24. Government NWP. The Bokone Bophirima (North West) Provincial Implementation Plan on HIV, TB and STIs (2017-2022). In: DOH, editor. North West2016. p. 5.
25. Services NHL. NHLS Strategic Plan 2010–2015. In: NHLS, editor. 2010. p. 1–50.
26. Goetz MB, Hoang T, Kan VL, Rimland D, Rodriguez-Barradas M. Development and validation of an algorithm to identify patients newly diagnosed with HIV infection from electronic health records. AIDS research and human retroviruses. 2014;30(7):626-33.
27. Kabudula CW, Clark BD, Gómez-Olivé FX, Tollman S, Menken J, Reniers G. The promise of record linkage for assessing the uptake of health services in resource constrained settings: a pilot study from South Africa. BMC Medical Research Methodology. 2014;14(1):71.
28. Sengayi M, Spoerri A, Egger M, Kielkowski D, Crankshaw T, Cloete C, et al. Record linkage to correct under-ascertainment of cancers in HIV cohorts: The Sinikithemba HIV clinic linkage project. International journal of cancer. 2016;139(6):1209-16.
29. J G. National Health Laboratory Service Corporate Data Warehouse patient linking technical document. 2014.
30. Pillay T, Cornell M, Fox MP, Euvrard J, Fatti G, Technau KG, Sipambo N, Prozesky H, Eley B, Tanser F, Johnson LF. Recording of HIV viral loads and viral suppression in South African patients receiving antiretroviral treatment: a multicentre cohort study. Antivir Ther. 2020;25(5):257-266. doi: 10.3851/IMP3371. PMID: 32960187; PMCID: PMC7982353.
31. Abdullahi, S.B., Ibrahim, O.R., Okeji, A.B. *et al.* Viral suppression among HIV-positive patients on antiretroviral therapy in northwestern Nigeria: an eleven-year review of tertiary care centre records, January 2009–December 2019. *BMC Infect Dis* **21**, 1031 (2021).
32. Pellowski JA. Barriers to care for rural people living with HIV: a review of domestic research and health care models. J Assoc Nurses AIDS Care. 2013 Sep-Oct;24(5):422-37.

doi: 10.1016/j.jana.2012.08.007. Epub 2013 Jan 22. PMID: 23352771; PMCID: PMC3640620.

33. Faber J, Fonseca LM. How sample size influences research outcomes. Dental Press J Orthod. 2014 Jul-Aug;19(4):27-9.
34. Chinyandura, C., Jiyane, A., Tsalong, X. *et al.* Supporting retention in HIV care through a holistic, patient-centred approach: a qualitative evaluation. *BMC Psychol* **10**, 17 (2022).
35. Steinbrink J, Imlay H, Rao K, Riddell J 4th. Identifying causes of persistent HIV viremia in adult patients at an academic medical center. SAGE Open Med. 2019
36. Palmer, A., Gabler, K., Rachlis, B., et al. Viral suppression and viral rebound among young adults living with HIV in Canada. Medicine 97(22):p e10562, June 2018.
37. Tilahun M, Gedefie A, Ebrahim E, Seid A, Ali A, Shibabaw A, Belete MA, Fiseha M, Tesfaye M, Ebrahim H, Abera A. Immuno-Haematological Abnormalities of HIV-Infected Patients Before and After Initiation of Highly Active Antiretroviral Therapy in the Antiretroviral Therapy Clinics of Six Health Facilities at Dessie Town, Northeast Ethiopia. J Blood Med. 2022 May 11;13:243-253.
38. Kahana SY, Jenkins RA, Bruce D, Fernandez MI, Hightow-Weidman LB, Bauermeister JA, et al. (2016) Structural Determinants of Antiretroviral Therapy Use, HIV Care Attendance, and Viral Suppression among Adolescents and Young Adults Living with HIV. PLoS ONE 11(4):

6.0 APPENDICES

Appendix 1.0: **Study explanatory variables**

Table 4: Study explanatory variables

Variable	Definition	Scale of measurement
Median Age	Health facility mean age of participants when the viral load measurement was done.	Numeric measured in years and is continuous.
Year	The year that the test was done	Categorical in nature by year e.g. 2011,2012 etc.
Median CD4 count	The median CD4 count for facility participants. This is a measure of CD4+ cells per milliliter.	Numeric continuous
Health facility geolocation	The GPS location of the various health facilities.	Consists of a latitude and a longitude, this is numerical and continuous.
Median Haemoglobin	The median facility hemoglobin. This is a protein that carries iron in the red cells.	It is a numerical and continuous in nature.
Median white cell count	The median facility white cell count. It is a measure of immune cells per unit volume of blood.	It is a numerical and continuous variable.

Data for these variables were summarized as median(IQR)

Appendix 2.0: Overview of National Health Laboratory Service Corporate Data Warehouse Patient Linking

Algorithm

The aim of the patient linking process is to identify specimens that were taken from the same individual and link them together by assigning a unique patient id to all these episodes. The linking process uses, as its basic principle, the following assumption: From the demographic attributes available from the laboratory database, an individual is distinguished from another based on the following attributes: (1) surname, (2) first name, and (3) birthdate.

The linking steps are as follows; the data cleaning and preparation stage, the exact linking step and the fuzzy linking step.

Data cleaning and preparation stage

The record received from the laboratory information system is cleaned and loaded to a raw demographics folder on the corporate data warehouse (CDW) system. The cleaning process includes trimming of blank spaces, removal of titles in name fields, removal of special characters, splitting of multiple given names in one variable into multiple variables, joining of surnames with prefixes into one name, converting all string variables to lower case and standardization of date formats.

Exact linkage stage

If the raw records have valid values for first name, surname and date of birth these attributes are matched against the values for existing unique patients in an archived folder with existing patients and, if matched, the unique patient identifier (id) number is assigned to the specimen. After this stage all raw records that have not been assigned a unique patient id number and have valid values for first name, surname and date of birth are assigned a new unique patient id number and records with the same first name, surname and date of birth are assigned the same unique patient id number. Lastly, all remaining raw records will be records for which insufficient data exists for any type of meaningful linking to be possible. Each of these will be assigned its own unique patient id as it is impossible for the process to determine if this relates to another specimen or not.

Fuzzy linkage stage

The fuzzy linking process uses as input the records generated by the last run on the exact linking process. The purpose of the fuzzy linking is to attempt to get a better level of linking by accommodating different spelling of names and surnames due to factors such as typing errors or complex spelling of surnames etc. In order to achieve this, rows of data are compared with each other and using the fuzzy algorithm to determine the probability of the same values from the 2 rows being the same, these probabilities are then weighted and the match is scored. Records that score above a certain cut-off level are then assumed to be the same and linked together. Since name and surname is considered in exact matching, exact matches on these attributes are excluded in fuzzy matching. The probability score is calculated for first name, surname and date_ of_ birth. The probability score is weighted as follows: First Name = 40%; Surname = 40% and Date of birth = 20%. If probability multiplied by weights is greater than 90% (cut-off level) then the record will be linked. All records that score a 90% or above match are extracted as matches to be resolved into a single hierarchy of linked demographics. This scoring and extraction of matched values results in a grid of unique patient's ids that are matched to each other. This is resolved into a single hierarchy resolving all values to a single unique patient id so that this id can be assigned to the other episodes to effectively link them together.

The detailed linking algorithm is available by request from the National Health Laboratory Service, South Africa. The patient linking process in its current release represents the best that has been achieved programmatically to date given the constraints involved such as volume of data, and especially the quality of data. It has been designed to favour linking in cases where concise attributes that should reasonably identify a patient, such as the national id number, is utilized. A conservative approach was used to underlink rather than overlink. However, due to factors such as the ability and limitations of the probability algorithm, both overlinking and underlinking is possible.

Appendix 3.0: Project timelines

Table 5: Project timelines

Year	2020				2021							2022										
Activity/ Month	**Sept**	**Oct**	**Nov**	**Dec**	**June**	**Jul**	**Aug**	**Sept**	**Oct**	**Nov**	**Dec**	**Jan**	**Feb**	**Mar**	**April**	**June**	**July**	**Aug**	**Sept**	**Oct**	**Nov**	**Dec**
Protocol corrections																						
Protocol Assessment																						
Ethics Submission																						
Protocol Approval																						
Data Retrieval																						
Data Cleaning																						
Data Analysis																						
Report Write up																						
Report Submission																						
Dissemination of Report																						

Appendix 4.0: Ethics Clearance Certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

Office of the Deputy Vice-Chancellor (Research and Innovation)

TO: Dr K Ndlovu
School of Public Health
Division of Epidemiology and Biostatistics
Medical School
University

E-mail: kudzynigel@gmail.com

CC: Supervisor: Dr K Otwombe
<otwombek@phru.co.za>
and <HREC-Medical Research Office@wits.ac.za>

FROM: Mr Iain Burns
Human Research Ethics Committee (Medical)
Tel: 011 717 1252

E-mail: Iain.Burns@wits.ac.za

DATE: 2021/08/19

REF: R14/49

PROTOCOL NO: **M210304** (This is your ethics application reference number. Please quote it in all enquiries, oral or written, relating to this study.)

PROJECT TITLE: *The spatial distribution of health facility viral suppression rates in the Dr Kenneth Kaunda District Municipality, South Africa, 2010-2015*

Please find attached the Clearance Certificate for the above project. I hope it goes well and that an article in a recognized publication comes out of it. This will reflect well on your professional standing and contribute to Government funding of the University.

MSWorks2000/Iain0007/Clearscan.wps

Figure 5: Ethics Clearance Certificate

Appendix 5.0: Data Access Request Letter

NATIONAL HEALTH LABORATORY SERVICE

Academic Affairs and Research
Modderfontein Road, Sandringham, 2031
Tel: +27 (0)11 386 6142
Fax: +27 (0)11 386 6296
Email: babatyi.kgokong@nhls.ac.za
Web: www.nhls.ac.za

24 June 2021

Applicant: Kudzayi Ndlovu
Institution: University of the Witwatersrand
Department: Epidemiology
Email: kudzynigel@gmail.com
Cell: 068 786 8862

CC: Kennedy Otwombe, Adrian Puren, Tendesayi Chakezha, Simbarashe Takuva

Re: Provisional Approval to access National Health Laboratory Service (NHLS) Data

Your application to undertake a research project **"The spatial distribution of health facility viral suppression rates in the Dr Kenneth Kaunda district municipality, South Africa, 2010-2015, Ref No: PR2010410"** using data from the NHLS database has been reviewed. This letter serves to advise that the application has been provisionally approved. For full approval to be granted and for you to be given access to the data you have requested, you need to satisfy the following conditions:

- Ethics approval is obtained from a recognised SA Health Research Ethics Committee.
- Processes are discussed with the relevant NHLS departments (i.e. Information Management Unit and Operations Office) and are agreed upon.
- Confidentiality is maintained at participant and institutional level and there is no disclosure of personal information or confidential information as described by the NHLS policy.
- A final report of the research study and any published paper resulting from this study are submitted and addressed to the NHLS Academic Affairs and Research office and the NHLS has been acknowledged appropriately.
- NHLS Data cannot be used to track patients as no pre-approval/consent is obtained from Patients.
- Tendesayi Chakezha is noted as NHLS collaborator for this research.

Please send any requested documents through to academic.research@nhls.ac.za. Once your application has received full final approval, we will confirm with you in writing. Should you wish to speak with us, please contact us on 011 555 0367.

Yours sincerely,

Dr Babatyi Malope-Kgokong
National Manager: AAR

Figure 6: Data Access Request

Appendix 6.0: Project management

The project timelines are summarized in the attached Gant-chart (Appendix 3.0), personal funds will be used to cover expenses at the various stages of the project.

Table 6: Budget for the study

Budget Line Item	**Unit Cost (R)**	**Number of Units**	**Total Cost (R)**	**Source of Funds**
Printing copies of Ethics Application form	30	10	300	Self-funded
Printing copies of Protocols	45	5	225	Self-funded
Printing of Research Report	90	5	450	Self-funded
Binding of Research Report	120	5	600	Self-funded
TOTAL	**R 285**		**R 1, 575**	

Appendix 7.0: Dissemination plan

Upon completion there will be a submission of findings to the University of Witwatersrand. We will pursue presentations at conferences within and outside South Africa. A submission will also be made to a peer reviewed journal for publication. Lastly since this study is based on program data, the findings will be disseminated to the North West province department of health, the NHLS and the health facilities in the study area This would present a unique opportunity for assessment of progress towards the 90-90-90 targets over the study period.