Politecnico di Torino, Université Paris Saclay

Quantifying the impact of internal migration on the risk of acquiring HIV in Namibia

Chiara Giaquinta

Supervisor: Dr. Eugenio Valdano

July 2023

Contents

1 Abstract

This thesis investigates the relationships between mobility, social/behavioral indicators, and HIV outcomes. We used data from the *2017 Namibia Population-Based HIV Impact Assessment (NAMPHIA)* and the *2011 Namibia Population Housing Census*. Focusing on specific sub-populations relevant to HIV transmission, the study examines how mobility patterns, along with various sociodemographic factors, influence HIV status and treatment. The findings reveal significant associations between age, social/behavioral factors, and HIV status and treatment. Moreover, while mobility exhibits a correlation with achieving Viral Load Suppression, no direct correlation with HIV status is observed. However, limitations arise from participants' hesitance to disclose sensitive information and the unavailability of more granular constituency-level mobility data. These findings contribute to the understanding of HIV dynamics in Namibia and provide insights for future interventions and policies aimed at reducing HIV transmission.

2 Introduction

Namibia has a generalized epidemic, with 12,6 % of the population living with HIV [1]. Despite the dramatic progress Namibia has made in HIV prevention and treatment, incidence is still too high in specific sociodemographic groups: 0,99 % of young women acquire HIV each year [1]. In the study of HIV transmission dynamics, migration is recognized as one of the key factors driving the virus diffusion, as supported by several studies [2]. Moreover, several studies have demonstrated that HIV prevalence is higher among migrants than among resident populations. For instance, a study conducted in 2008 in Kilimanjaro revealed that rural in-migrants were at a higher risk of contracting HIV, and this significantly contributed to the increased prevalence of HIV in rural communities [3]. Another study carried out in Uganda in the early 1990s showed that HIV prevalence was 16% among those who had migrated in the previous 3 years, compared to 6% among non-migrants [4].

In particular, Namibia is characterized by a highly mobile population [5] and a high HIV prevalence. The population, amounting approximately to 2.5 million people, is widely dispersed. Processes of urbanization, poverty, and extreme climatic events caused and continue to cause significant migration flows, both over long and short distances. Another main cause of migration is linked to employment opportunities, particularly among young adults. Namibia's economy heavily relies on farming, fishing, mining and tourism, creating a labor market that attracts both men and women seeking employment [6]. On the administrative level, in 2013 Namibia is divided into 14 regions, each of them further divided into constituencies, with a total of 121 constituencies. It shares borders with Angola and Zambia to the north, Botswana and Zimbabwe to the east and South Africa to the south, while the Atlantic Ocean represents its western border [7].

Namibia's unique demographic and geographic characteristics, coupled with its economic activities and migration patterns, create a complex context for the study of the relationship between population mobility and the spread of HIV. The main objective of this thesis is the investigation of potential correlations between mobility and HIV-related indicators, such as HIV status and Viral Load Suppression (VLS)[1]. Previous studies in Namibia have used Call Detailed Records (CDRs) from mobile phones [8] to examine these relationships. These studies have demonstrated that the 40% of the infection risk is linked to mobility, highliting the significant role of mobility-driven networks in the HIV epidemic in Namibia. Instead, in this project data collected from *2017 Namibia Population-based HIV Impact Assessment (NAMPHIA)* [9], containing information on both migration and HIV, have been analyzed [2]. NAMPHIA represents the first comprehensive national survey in Namibia collecting data on both migration and HIV-related information. Furthermore, to have a more complete understanding of the socio-demographic features of Namibian population, microdata from *2011 Namibia Population Housing Census* [10] have also been explored and analyzed. Census data serves as the main source for demographic and socio-economic statistics available for Namibia, since it provides information on the size, distribution, composition and several social and economic characteristics of the population.

In Section 3, the methodolody followed to study potential correlations between variables of interest is presented, together with its mathematical formulation. In Section 4, results are shown on both the correlation analysis and on the examination of the socio-demographic profile of the Namibian population.

The ultimate goal of this project is to understand if specific groups of people may be more vulnerable in acquiring HIV or less likely to be under treatment, due to their mobility. In the fight against the spread of the virus, these findings could play a crucial role in the development of new strategies specifically tailored to target certain sub-populations. By doing so, resources for HIV prevention and treatment can be deployed more effectively and efficiently, maximizing their impact in fighting the spread of the virus.

[1]Effective HIV treatment aims to suppress the replication of the virus in the body and maintain optimal health for individuals living with HIV. This treatment typically involves the use of antiretroviral therapy (ART) medications. One important indicator of successful treatment is the achievement of Viral Load Suppression (VLS), which is defined as having less than 1,000 HIV ribonucleic acid (RNA) copies per milliliter (mL) in the plasma. VLS indicates that the level of HIV in the blood is effectively controlled, minimizing the risk of disease progression and reducing the likelihood of transmitting the virus to others.

[2]NAMPHIA is part of the PHIA Poject, implemented by ICAP at Columbia University in partnership with the Ministries of Health and the US Centers for Disease Control and Prevention (CDC). It is designed to measure the impact of HIV programs in PEPFAR-support countries through national household surveys [9].

3 Methods

3.1 Logistic regression model

A Bayesian inference approach has been adopted, together with MCMC (Monte Carlo Markov Chain) sampling. This Bayesian framework enables the analysis of associations between social and behavioural variables and HIV-related outcomes, providing meaningful insights into the relationships between them.
In Bayesian inference, the goal is to estimate the posterior distribution P($\theta|d$), which represents the statistical distribution of the model parameters θ given the observed data d. According to Bayes' theorem, the posterior distribution is proportional to the product of the likelihood function P($d|\theta$) and the prior distribution P(θ), divided by the marginal distribution P(d):

$$P(\theta|d) = \frac{P(d|\theta)P(\theta)}{P(d)} \tag{1}$$

Since obtaining the exact form of the posterior distribution is often analytically intractable, Markov Chain Monte Carlo (MCMC) methods are used to sample from the posterior distribution, allowing for efficient exploration of the parameter space. By sampling from the posterior distribution using MCMC, a large number of parameter values are generated, which can be used to estimate various posterior statistics of interest, such as the median and the credibility intervals.
In the specific case of the univariate logistic regression, if y is the binary outcome and p its associated probability [3], x is the indipendent variable and α and β are the model parameters, then the Bernoulli likelihood can be expressed as follows:

$$P(x, y|\alpha, \beta) = py + (1-p)(1-y) = \begin{cases} p & \text{if y=1} \\ 1-p & \text{if y=0} \end{cases} \tag{2}$$

The parameter p of the Bernoulli likelihood is parametrized using a logistic link function μ, which relates the independent variable x to the probability p of the binary outcome y:

$$\mu = log\frac{p}{1-p} = \alpha + \beta x \tag{3}$$

As such, the logistic regression belongs to the family of generalized linear models, since the link function allows for a linear relationship between x and the logarithm of the odds ratio of y. This enables the modeling of the probability p as a function of x.

Considering all the N individuals in the targeted population, the full likelihood can be expressed as the product of individual likelihood terms:

$$L = \prod_{i=1}^{N}[p_i y_i + (1-p_i)(1-y_i)] \tag{4}$$

[3]It has been assumed that the binary outcome variable $y \in$ 0,1 has a Bernoulli likelihood

Here, p_i and y_i refers to the i-th individual.
In a multilevel model, also the region of residence r can be taken into account. This introduces a new, regional level of analysis, and hence an intercept likelihood for each region r, $\alpha_r \sim Normal(\mu^{(\alpha)}, \sigma^{(\alpha)})$. The introduction of α_r and its normal distribution allows capturing the variability of intercepts among regions, enabling the assessment of the specific effect of the region of residence on the considered outcome. The full likelihood becomes:

$$L = \prod_{r=1}^{14} N(\alpha_r | \mu^{(\alpha)}, \sigma^{(\alpha)}) \prod_{i=1}^{N} [p_i y_i + (1 - p_i)(1 - y_i)] \tag{5}$$

Here, the logistic link function becomes:

$$\mu_i = log \frac{p_i}{1 - p_i} = \alpha_{r_i} + \beta x_i \tag{6}$$

where r_i represents the region of the i-th individual.

3.2 Influence of the regions of past residence

The multivariate logistic regression model can be enhanced by considering also the influence of the regions of past residence for each individual. The introduction of a new parameter $\gamma \in [0, +\infty[$, which represents the weight of the previous regions of residence, allows for capturing their effect. Meanwhile, the contribution of the current region is set to 1. Thus, the logistic link in Equation 6 can be reformulated as follows:

$$\mu_i = log \frac{p_i}{1 - p_i} = A_i + (\beta x)_i \tag{7}$$

Indicating with α_r the coefficient associated to region r and assuming that the region of current residence is denoted by c, the coefficient A_i is given by:

$$A_i = \epsilon \left(\alpha_c + \gamma \sum_{k=1, k \neq c}^{K<14} \alpha_k \right) \tag{8}$$

and the summation is performed over the subgroup K of regions where the i-th individual previously resided. The coefficient ϵ is a normalization factor which could be determined considering the special situation where all the regions have the same effect, so $\alpha_r = \mu_\alpha \ \forall r$. Then, the coefficient A_i becomes:

$$A_i = \epsilon \mu_\alpha (1 + K\gamma) \tag{9}$$

Setting $\epsilon = \frac{1}{K\gamma}$, Equation 8 can be rewritten as:

$$A_i = \frac{\left(\alpha_c + \gamma \sum_{k=1, k \neq c}^{K<14} \alpha_k \right)}{1 + K\gamma} \tag{10}$$

It emerges that the coefficient A_i is simply the weighted average of the contribution of the regions in which the individual has resided.

It is possible to make a further generalization. Let's consider the matrices Z and W, whose elements are given respectively by:

$$Z_{ir} = \begin{cases} 1 & \text{if the i-th individual currently lives in region r} \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

$$W_{ir} = \begin{cases} 1 & \text{if the i-th individual used to live in region r (but not anymore)} \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Additionally, it is possible to define the array V_i as the sum of the elements in the i-th row of the W matrix, representing the number of regions where the individual lived throughout his life in the past:

$$V_i = \sum_r W_{ir} \tag{13}$$

Finally, the Equation 10 becomes:

$$A_i = \frac{\sum_r (Z_{ir} + \gamma W_{ir})\, \alpha_r}{1 + \gamma V_i} \tag{14}$$

3.3 Data processing for correlation analysis

The Bayesian approach introduced in Paragraphs 3.1 and 3.2 has been applied to investigate potential associations between social and behavioural variables and HIV-related indicators, such as the Viral Load Suppression and the HIV status. Given that the HIV landscape for young women in Namibia and Southern Africa has already been extensively explored [14] [15] [16], the focus of this analysis has been on the male population aged 15-45 years. This particular group exhibits higher mobility and higher sexual activity, thus potentially representing the main vector for the virus transmission to young women. Indeed, it is important to consider that in sub-Saharan Africa, there is a tendency among young women to engage in relationships with older men, often involving material exchanges. [11].

For this category, a set of variables has been carefully selected, to investigate their possible influence on HIV-related outcomes. Their selection was based on the findings from the social and economic analysis which will be presented in Section 4.1, as well as a comprehensive review of relevant literature [12] [17] [18].

Whenever possible, a binarization of these variables have been performed. In Table 1, the variables of interest, together with their numerical values, have been reported.

Table 1: Definition of variables

Variable name	Values
Age	*int* from 15 up to 45
Mobility[a]	*int* from 0 up to 13
Education level[b]	0 if *missing*, 1 if *primary*, 2 if *secondary or higher*
Economic status	0 if *poor*, 1 if *wealthy or rich*
Area	0 if *urban*, 1 if *rural*
Employment status	0 if *unemployed* and 1 if *employed*
Alcohol consume	0 if *teetotal*, 1 if *alcohol consumer*

[a] It corresponds to the number of regions of past residence (the region of current residence is excluded).

[b] Many respondents did not provide an answer to this question in the survey. However, in order take this data into account, the possible outcome *missing* has been introduced.

4 Results

In this section, results obtained through the analysis of data from *2011 Namibia Population Housing Census* and from *2017 Namphia Population-based HIV Impact Assessment (NAMPHIA)* are presented. The exploration of these datasets, detailed in Paragraph 4.1, allowed to examine various factors, including HIV-related variables, mobility patterns, and behavioral and social factors. Furthermore, using the dataset provided by NAMPHIA, the methods outlined in Section 3 were applied specifically to the sub-population of men aged between 15 and 45 years, and the corresponding results are presented in Paragraph 4.2.

4.1 Migration and HIV: data analysis

To gain a deeper understanding of the migration patterns within the more mobile sub-populations, an in-depth analysis of microdata from *2011 Namibia Population Housing Census* has been conducted[4].
The population of Namibia amounted to 2.113.077 people, with women accounting for 51,6 % and men representing the remaining 48,4 %. The age distribution of the entire population is shown in the histogram on the top of the Fig. 1. Even if the female population is larger than the male population, men tend to exhibit higher mobility than women, as shown in the bar plot on the bottom of the Fig. 1. Notably, the most mobile sub-population consists of young men and women aged between 20 and 29, a period that coincides with their peak sexual activity. Approximately the 7,8 % of men and the 6,9 % of women in this age bracket

[4] In particular, to derive 1-year internal migration statistics, the census (with a reference night of 28 August 2011) asked about each individual's place of usual residence (where a person usually lives) and the place of previous residence (where the person lived since September 2010). An individual is considered a migrant if the constituency of usual residence and the constituency of previous residence do not match. It must be remembered that at the time of 2011 Census Namibia had 13 regions and 107 constituencies.

changed residence.

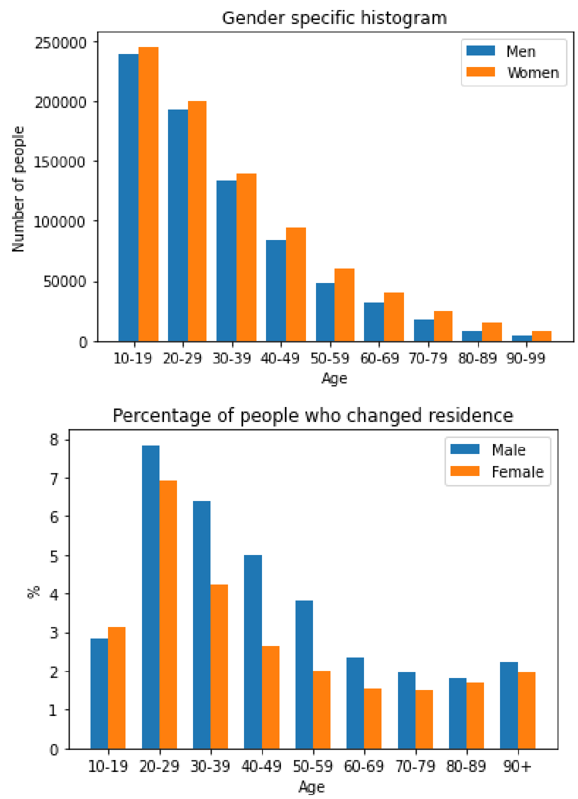

Figure 1: On the top, the gender and age specific histogram for the whole population. On the bottom, the bar plot for the respondents who changed residence. This bar plot represents the percentage of men and women who moved to a different constituency, with respect to the total male and female population in the same age bracket.

Moreover, it has been also possible to examine the contribution of international migrants, specifically those who had the previous residence outside Namibia. They amount approximately to the 1 % of the entire population. In particular, the 76,5 % of the international immigrants consists in African immigrants, especially from Angola, South Africa, Zimbabwe, Zambia and Botswana. Considering the HIV prevalence data associated with these specific countries [13], it is possible to gain insights into the extent of their impact on the spread of HIV in Namibia. To visualize this relationship, refer to the scatter plot in Fig. 2.

To identify constituencies that are undergoing population growth or decline, a flux map (refer to Fig. 3) has been generated for both sexes, focusing specifically on individuals between the ages of 15 and 30 who have never been married.

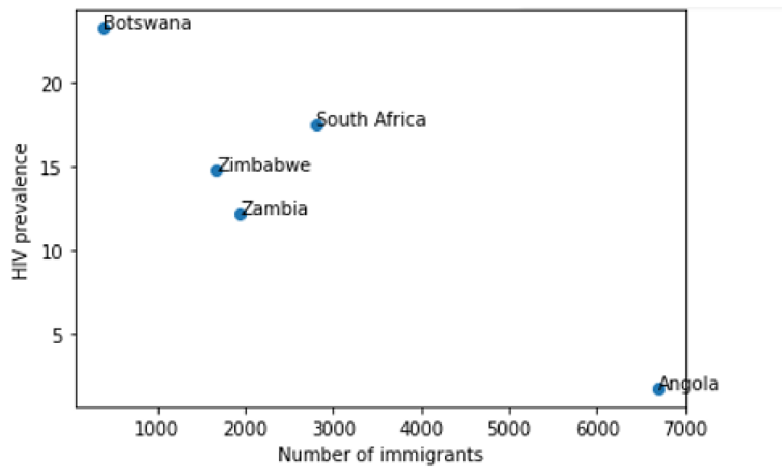

Figure 2: Scatter plot with HIV prevalence on the y-axis and the number of immigrants on the x-axis; the focus is just on specific African immigrants.

This category has been chosen as it is considered to be the most sexually active and, therefore, potentially the most influential in terms of HIV transmission.

The flux maps provide valuable information about population dynamics in each constituency. The flux, represented by *f*, is calculated using the equation:

$$f = \frac{N_{in} - N_{out}}{\frac{(N_{stay}+N_{out})+(N_{stay}+N_{in})}{2}} \tag{15}$$

Here, N_{in} represents the number of people entering the constituency, N_{out} the number of people leaving the constituency and N_{stay} the number of people who remained in the constituency.
Additionally, churn maps (Fig. 4) provide further insights into the turnover within each constituency, for the target sub-population of never married women and men aged 15-30 years. These maps are a useful tool to understand the level of involvement of a constituency in migration flows.

As in the flux maps, also in the churn maps a value c is associated to each constituency, and it is given by:

$$f = \frac{N_{in} + N_{out}}{\frac{(N_{stay}+N_{out})+(N_{stay}+N_{in})}{2}} \tag{16}$$

Where N_{in}, N_{out} and N_{stay} have the same definitions as in Equation 15. These insights could provide a deeper understanding of the most significant migratory patterns. It is notable that the constituencies in the Khomas region, particularly those of Windhoek (the capital), exhibit the highest influxes and population turnovers.

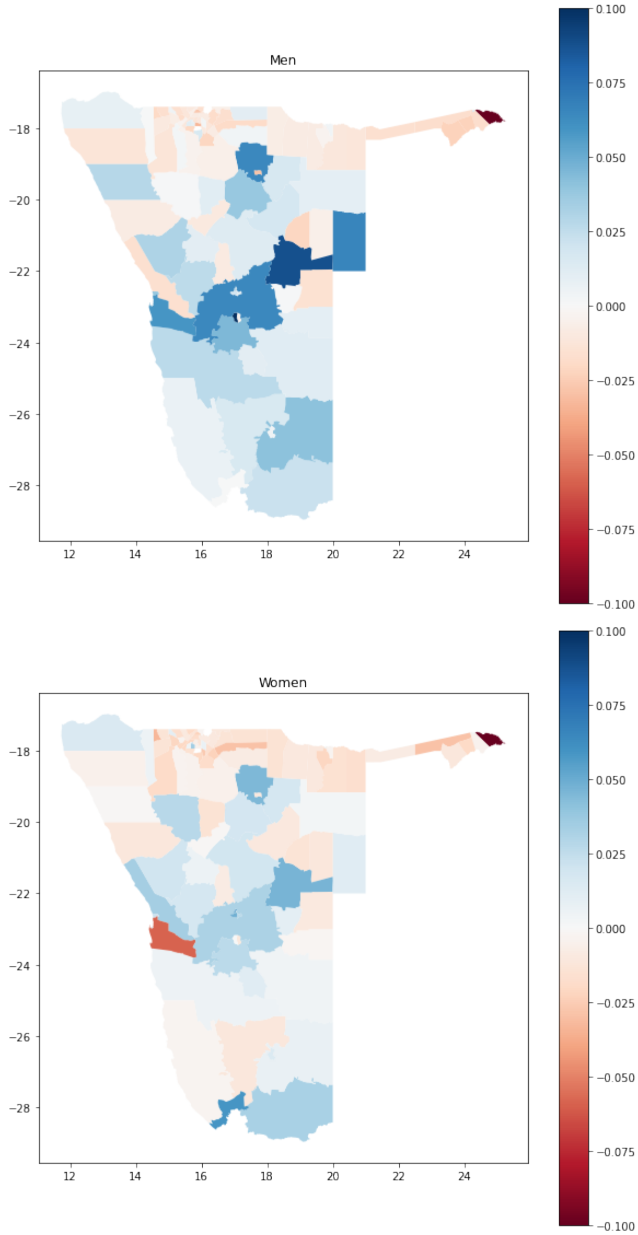

Figure 3: Flux maps for never married men (on the top) and women (on the bottom), with age between 15 and 30. Constituencies in red are those experiencing population decrease (f ¡ 0), constituencies in blue are those experiencing population increase (f ¿ 0) and constituencies in white show no change in population (f = 0). The color scale corresponds to the proportion of never married men/women aged 15-30 who changed their residence, normalized to the same sub-population within each specific constituency. Notably, the Kabbe Constituency exhibits a significant decrease, which is likely attributed to a flood occurred in 2010 [5].

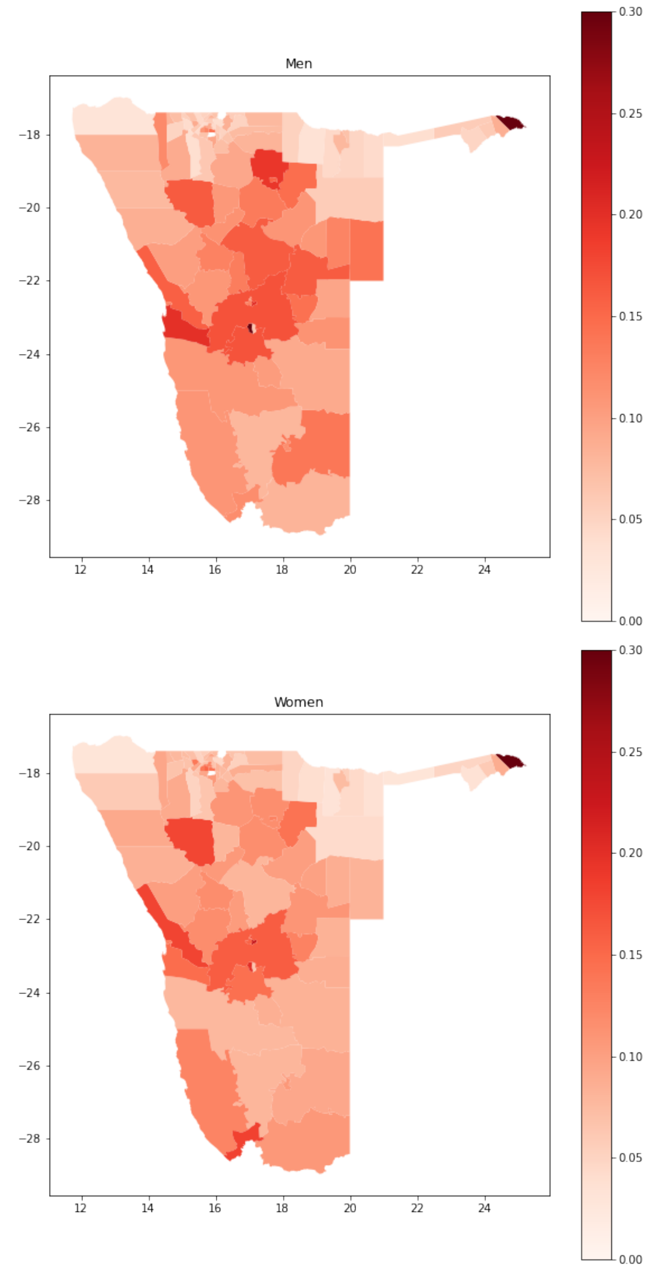

Figure 4: Churn maps for never married men (on the top) and women (on the bottom), with age between 15 and 30 years.

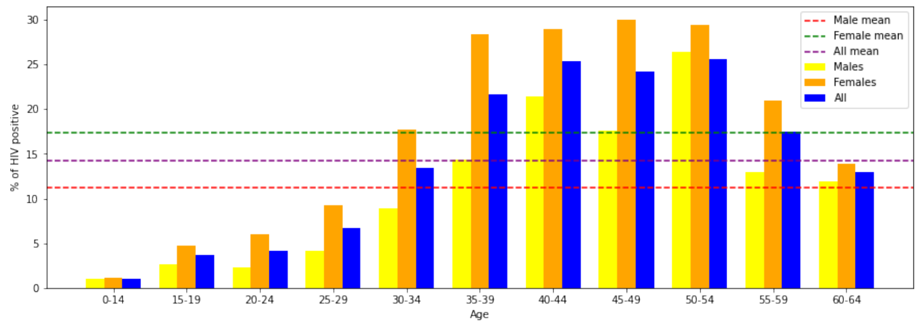

Figure 5: Bar plot of the national HIV prevalence, for men, women and both sexes, together with their average across ages.

Exploring the data collected from *2017 Namphia Population-based HIV Impact Assessment (NAMPHIA)*, a general overview of the HIV situation in Namibia in 2017 has been obtained [5]. In particular, the bar plot in Fig. 5 presents the estimated national HIV prevalence[6] by age; higher prevalence rates were found among women aged 45-49 (30%) and men aged 50-54 (26.4%).

Additionally, the data shows a significant finding regarding the awareness of HIV status among young respondents. Specifically, the survey indicates that 56.6% of men and 37.4% of women aged 15-24 years were unaware of their HIV-positive status at the time of the survey. This information becomes particularly relevant considering the high migration rates among young men, as it sheds light on the potential for HIV transmission due to this subgroup. Understanding the dynamics of the awareness of HIV status and migration patterns can be instrumental in comprehending the spread of the HIV epidemic.
To gain additional insights, Fig. 6 presents a bar plot illustrating the national Viral Load Suppression (VLS) rates.

At the regional level, an analysis on the VLS, prevalence and awareness of self HIV-status has been conducted. In Fig. 7, the HIV prevalence and the VLS rates for the population aged 15-64 years are shown at the regional level. It is evident that although the northern regions have a higher prevalence, they also have a higher proportion of individuals achieving VLS.

Combining the information on VLS and HIV prevalence at a regional level

[5]This cross-sectional survey exploited a two-stage sampling design to obtain a nationally representative sample of adults aged 15-64 years and of children aged 0-14 years, living in households in Namibia in 2017. Based on the *2011 Population and Housing Census* [10], the sampling frame included all households in the country. In the second stage of the sampling process, households were randomly selected from each Enumeration Area (EA). Consenting heads of households completed a household questionnaire, including a roster of all residents and guests. People in a selected household were eligible if they had slept in the house the night before. Eligible individuals for each household were asked to consent to a questionnaire including sociodemographic and behavioral risk questions and home-based HIV testing.

[6]HIV prevalence refers to the percentage of people in a population who are infected with HIV at a particular time.

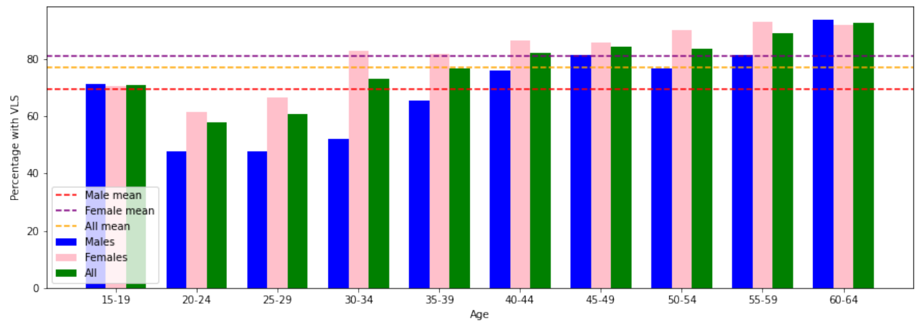

Figure 6: Bar plot with the national VLS, for men, women and both sexes, together with their average across ages.

can provide valuable insights. The result is shown in fig. 8.

In this case, a value k has been assigned to each region. Its value is given by:

$$k = p(1 - t) \tag{17}$$

where p represents the HIV prevalence and t represents the VLS.

Regarding HIV status awareness, it emerges that women generally have higher levels of awareness than men. Specifically, regions with lower HIV status awareness among men are are Kunene, Khomas, Otjozondjupa, Omaheke, and Karas. On the other hand, Karas, Hardap, and Omaheke are the regions where the percentage of men aware of their HIV-positive status but not receiving treatment is higher.

As previous studies have shown [12], certain occupational categories in Namibia are associated with a higher HIV prevalence. In particular, among the most common female professions, domestic workers and manufacturing workers appear to be associated with a higher prevalence. Among men, on the other hand, fishermen and truck drivers are subject to a higher prevalence. Therefore, using the detailed information on respondents' main activities from the *2011 Namibia Population Housing Census*, an analysis of professional profiles was conducted to investigate if the occupations at higher risk of HIV infection also tend to have higher mobility rates. The focus was to understand whether individuals in these occupational categories are more likely to migrate.

Initially, the proportion of employed individuals, unemployed individuals and students was examined. In particular, only people who provided an answer to the corresponding question in the survey (amounting to the 78,8 % of the total number of people in the dataset), and with age between 15 and 54 years old, have been taken into account. It has been found that 40 % were employed, 26.1 % were unemployed, and 17.1 % were students[7]. It is interesting to examine

[7] However, it should be noted that the available data on male and female students above the age of 25 are limited, making it difficult to obtain reliable results about the mobility of

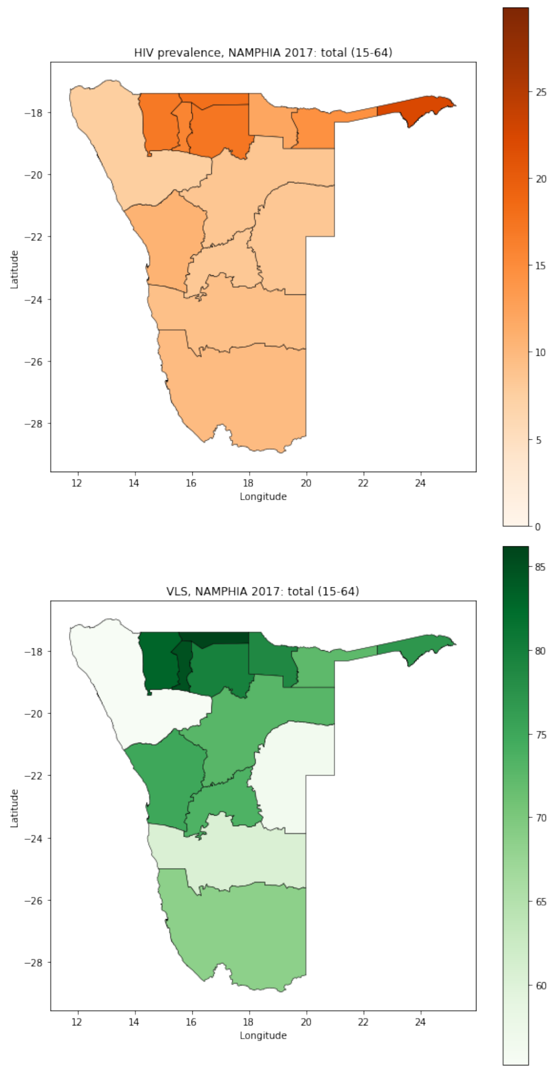

Figure 7: HIV prevalence (top) and VLS (bottom) at the regional level, for the population aged 15-64 yers.

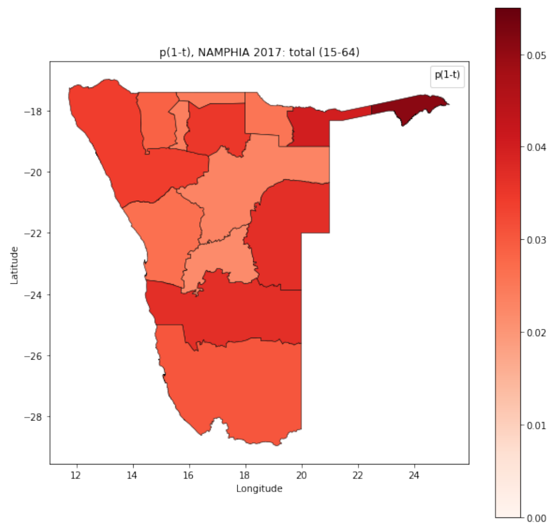

Figure 8: Map which combines the information on HIV prevalence (p) and on the VLS (t). The map refers to the population aged 15-64 years.

migration rates for these categories. The results, depicted separately for men and women, are presented in Fig. 9.

This analysis reveals distinct migration rates according to the age bracket; in particular, younger employed individuals exhibit higher migration rates compared to individuals in the same age bracket.

A further analysis can be conducted by considering the specific professions of employed individuals in the dataset, focusing on the age group between 15 and 54 years[8]. Specific job categories have been selected for men and women, taking into account the known differences in their employment sectors [12]. The results of this analysis are presented in Figure 10.

Among female workers, it is observed that young domestic workers and young women working in the manufacturing industry have higher migration rates. Among all young male workers, mobility is generally high, with the exception of those working in the primary sector, which represents the leading employment sector in the whole country [6].
Finally, it is also possible to study how mobility varies based on the highest level of education among respondents aged 15 to 54 years old. The results are reported in Fig. 11.

The analysis reveals that college students, regardless of their gender, exhibit higher migration rates compared to individuals in the same age groups. In particular, it has been noted that there is a significant influx students with

this specific subgroups.

[8]However, it must be remarked that only the 26,1 % of the individuals in the original dataset provided an answer to this question.

Figure 9: Bar plots representing the percentage of men and women in each category (employed,unemployed or student) who changed residence, compared to men and women in the same category and age brackets. Categories are distinguished by using different colors. Red dots represent mobility rates for the entire original dataset. Blue represent mobility rates for the modified dataset, which includes only individuals who provided a response to the question in the survey.

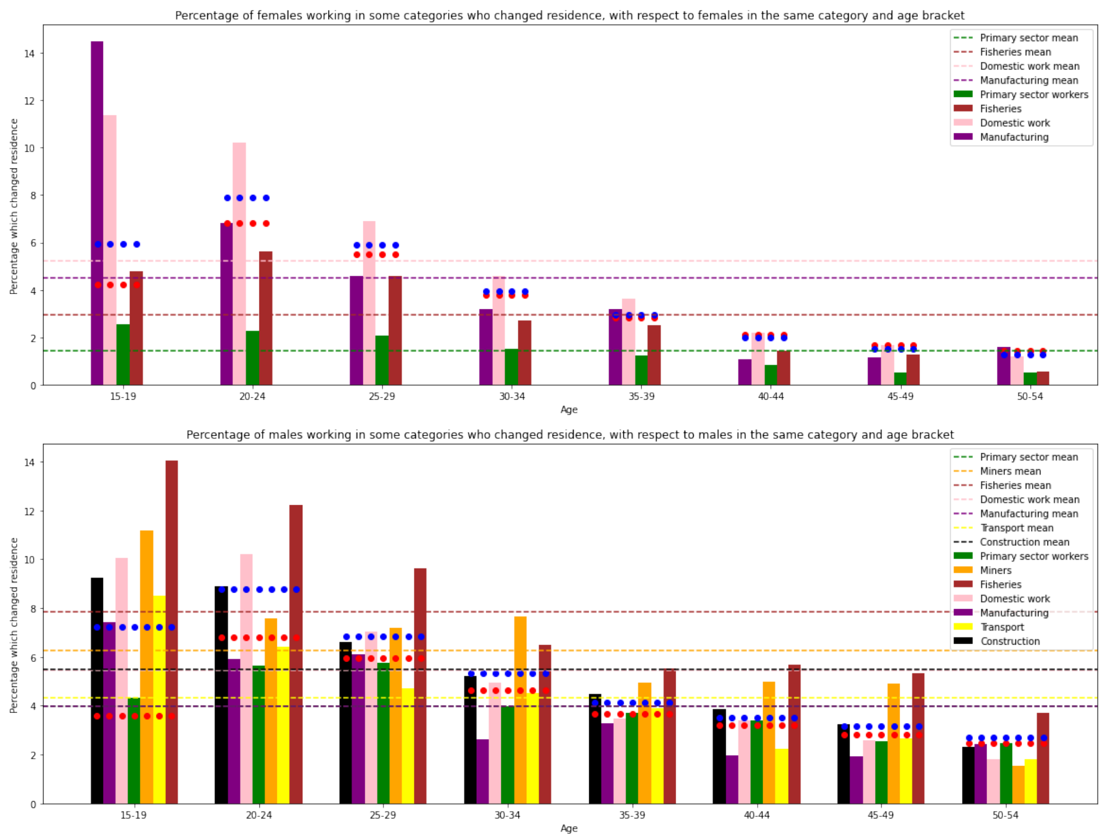

Figure 10: Bar plots representing the percentage of men or women employed in specific sectors who changed residence, compared to men or women in the same category and age brackets. Categories are defined using different colors. Red dots represent mobility rates for the entire original dataset. Blue represent mobility rates for the modified dataset, which includes only individuals who provided a response to the question in the survey.

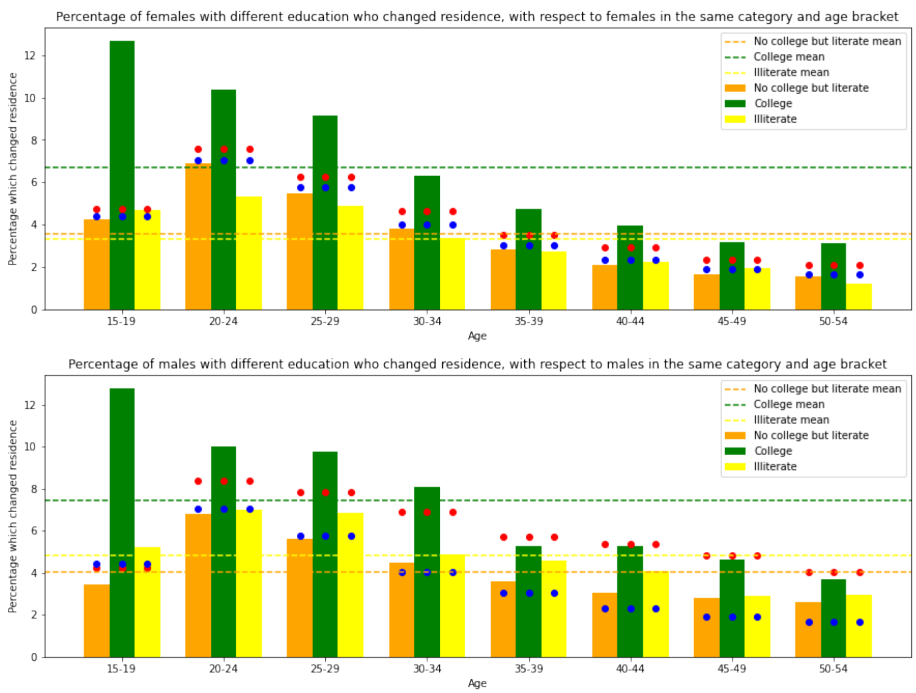

Figure 11: Bar plots representing the percentage of female and male students who changed residence, with respect to female and male students in the same age brackets. Categories are defined using different colors. Red dots represent mobility rates for the entire original dataset. Blue represent mobility rates for the modified dataset, which includes only individuals who provided a response to the question in the survey.

higher education towards the constituencies in the Khomas region, where the capital city of Windhoek is located.

4.2 Statistical inference: determinants of HIV status

For this analysis, two different models have been adopted. The first one (*Model 1*) is based on the theory presented in Paragraph 3.1, where the influence of the regions of past residence was not taken into account. The second one instead (*Model 2*) is based on the theory presented in Paragraph 3.2, and considers also the influence of the regions of past residence.

Two possible outcomes have been chosen: HIV final status and VLS.

1) **First outcome: HIV final status**

Initially, HIV final status has been considered as the output y of the model[9]. In particular:

$$y = \begin{cases} 1 & \text{if the respondent is HIV-positive} \\ 0 & \text{if the respondent is HIV-negative} \end{cases} \tag{18}$$

The two models have been then implemented for the same output. Regarding *Model 1*, different combinations of features have been explored and the odd ratios are collected in Table 2, along with the corresponding 95 % credibility intervals [10].

Table 2: Model 1: Odd Ratios median and 95 % Credibility Intervals for different combinations tested

Variables Combination	Age	Mobility	Employment status	Alcohol consumption	Economic status
Age	1.11 [1.10,1.12]				
Age + Mobility	1.11 [1.10,1.12]	0.99 [0.89,1.10]			
Age + Employment status	1.11 [1.10,1.13]		0.55 [0.44,0.68]		
Age + Alcohol consumption	1.11 [1.10,1.12]			0.62 [0.50,0.76]	
Age + Economic status	1.11 [1.10,1.12]				0.75 [0.60,0.92]

[9] It must be remembered that this analysis has been applied to the 82,5 % of the original dataset, for which it has been possible to establish the HIV final status.

[10] Only the variables characterized by credibility intervals not consistent with 1 have been presented here, with the only exception of the variable *mobility*.

	Age	Mobility	Employment status	Alcohol consumption	Economic status
Age + Economic status + Alcohol consumption + Employment status	1.12 [1.11,1.13]		0.58 [0.47,0.72]	0.63 [0.52,0.78]	0.80 [0.64,1.00]

Also in the case of the application of *Model 2*, several combinations of features have been considered, and odd ratios are collected in Table 3, together with the corresponding 95 % credibility intervals[11].

Table 3: Model 2: Odd Ratios median and 95 % Credibility Intervals for different combinations tested

Variables Combination	Age	Mobility	Employment status	Alcohol consumption	Economic status
Age	1.11 [1.10,1.12]				
Age + Mobility	1.11 [1.10,1.12]	1.03 [0.93,1.14]			
Age + Employment status	1.12 [1.10,1.13]		0.57 [0.46,0.71]		
Age + Alcohol consumption	1.11 [1.10,1.13]			0.62 [0.51,0.76]	
Age + Economic status	1.11 [1.10,1.12]				0.75 [0.60,0.92]
Age + Economic status + Alcohol consumption + Employment status	1.12 [1.11,1.13]		0.60 [0.48,0.75]	0.63 [0.51,0.78]	0.83 [0.67,1.03]

2) **Second outcome: VLS**

In the second step of the analysis, also the Viral Load Suppression has been considered as a possible outcome y[12]. In this case:

[11] Also in this case, only the variables characterized by Credibility Intervals not consistent with 1 have been presented here, with the only exception of the variable *mobility*.

[12] It must be remembered that in this case the number of respondents in the dataset has been drastically reduced, since only the 7,1 % of the original dataset has got a defined final VLS status.

$$y = \begin{cases} 1 & \text{if the respondent has reached VLS} \\ 0 & \text{if the respondent has not reached VLS} \end{cases} \quad (19)$$

Only *Model 1* has been used in this case[13]. Different combinations of the variables have been explored, and only the odd ratios whose credibility intervals are not consistent with 1 are collected in Table 4.

Table 4: Model 1: Odd Ratios median and 95 % Credibility Intervals for different combinations tested

Variables Combination	Age	Mobility	Employment status	Alcohol consumption
Age	1.05 [1.02,1.07]			
Age + Mobility	1.05 [1.03,1.08]	0.78 [0.64,0.93]		
Age + Employment status	1.06 [1.03,1.09]		0.38 [0.25,0.59]	
Age + Alcohol consumption	1.06 [1.04,1.09]			0.25 [0.16,0.39]
Age + Mobility + Employment status	1.06 [1.04,1.09]	0.80 [0.65,0.97]	0.40 [0.25,0.61]	
Age + Mobility + Alcohol consumption	1.07 [1.04,1.10]	0.74 [0.61,0.89]		0.24 [0.15,0.38]
Age + Mobility + Alcohol consumption + Employment status	1.08 [1.05,1.11]	0.76 [0.63,0.92]	0.42 [0.26,0.68]	0.24 [0.15,0.39]

5 Conclusions

The analysis of the census data revealed that young adults aged 20-29, who are also at their peak sexual activity, exhibit high levels of mobility; approximately 7.8 % of men and 6.9 % of women in this age group changed their residence. Furthermore, the analysis explored the contribution of international migrants.

[13]In reality, also *Model 2* has been tested to fit the data. However, the results are not shown, as the Credibility Intervals for the parameter γ were excessively wide, and this could lead to less precise estimates.

In particular, migrants from Angola, South Africa, Zimbabwe, Zambia, and Botswana account for the majority (76.5 %) of international immigrants. By considering the HIV prevalence in these countries, insights into their impact on HIV spread in Namibia can be gained.

Flux maps were generated to identify constituencies experiencing population growth or decline, focusing on never-married individuals aged 15-30, who are considered the most sexually active and influential in HIV transmission. Churn maps were also examined to understand the turnover within each constituency for this target sub-population. The Khomas region, particularly the capital Windhoek, exhibited the highest influxes and population turnovers.

The analysis of the data from NAMPHIA data provided a general overview of the HIV situation in Namibia in 2017. Higher HIV prevalence rates were found among women aged 45-49 and men aged 50-54. Notably, a significant percentage of young respondents, particularly men, were unaware of their HIV-positive status at the time of the survey. This finding is crucial considering the high migration rates among young men and their potential contribution to HIV transmission.

Viral Load Suppression rates were examined at the national level, revealing variations between men and women. Regional analysis showed that the northern regions, despite having higher prevalence, also had a higher proportion of individuals achieving VLS. Combining HIV prevalence and VLS information at the regional level provided further insights into the dynamics of HIV in Namibia.

Occupational profiles were analyzed to investigate if occupations associated with higher HIV prevalence also had higher mobility rates. Migration rates for different categories were explored, with men and women showing different patterns. In particular, high mobility rates for men were found among fishermen, truck drivers and miners. For women, high mobility rates were found among domestic workers and manufacturing workers. Also college students, both males and females, exhibit high mobility rates.

Overall, the analysis of these datasets provided valuable insights into migration patterns, HIV prevalence, awareness, and related factors in Namibia. These findings contribute to a better understanding of the spread of the HIV epidemic and can inform targeted interventions and prevention strategies.

Regarding the analysis of correlations between social and behavioral variables, mobility, and HIV-related quantities, only men between the ages of 15 and 45 were considered. Considering the final VLS status as the outcome, there is a clear association with mobility. Individuals who have resided in different regions throughout their lifetime exhibit a lower likelihood of achieving VLS compared to those who have remained in the same region consistently. This observation suggests that mobility may have an impact on the ability to achieve Viral Load Suppression. Further research is necessary to investigate this specific mechanism. One possible explanation is the disruption of healthcare continuity that can occur when individuals relocate. Additionally, social and support networks play a crucial role in managing HIV care. Relocating to a different region may result in a loss of social support systems, which can have negative implications for treatment adherence.

A positive association has been identified with age, indicating that older men are more likely to achieve VLS. This is mostly because they are more likely to be aware of their status. Furthermore, an association has been observed also with alcohol consumption, as people who do not consume alcohol exhibit a higher likelihood of attaining VLS. A peculiar finding, however, is the correlation between employment status and VLS. Surprisingly, employed individuals are less likely to achieve Viral Load Suppression compared to those who are unemployed. This finding is in contrast with what other studies have shown [19]. One possible explanation for this unexpected relationship could be the influence of socioeconomic factors. It is plausible that individuals who are unemployed may have better access to healthcare resources, including consistent medical appointments and adherence to treatment, due to their potentially lower income levels and eligibility for social welfare programs. On the other hand, employed individuals may face additional challenges in maintaining regular medical visits and adhering to treatment due to work-related responsibilities and time constraints. However, it is important to note that these findings are preliminary, and further research is needed to delve into the underlying mechanisms and factors contributing to this association.

With respect to the final HIV status as the outcome, no significant correlations with mobility were instead observed, consistent with findings from previous studies [12].

Furthermore, it is evident that this specific demographic subgroup exhibits minimal dependence on residential location (rural/urban) and educational attainment. However, a positive correlation with age emerges, indicating that older men are more susceptible to be HIV-positive.

Additionally, employment status and economic well-being show a significant association with the final HIV status; specifically, employed individuals and those belonging to the medium-high wealth quintile display lower HIV prevalence.

Interestingly, a noteworthy correlation is observed between alcohol consumption and HIV infection, wherein people who do not consume alcohol exhibit a higher likelihood of being HIV-negative. To provide a plausible justification for this observation, additional research is necessary. It is possible that individuals who abstain from alcohol may engage in other high-risk behaviors or have different patterns of social interaction that contribute to their increased susceptibility to HIV.

This project presents some limitations. Firstly, some participants chose not to respond to certain questions on their sexual behaviour, possibly due to social stigma. This restricted the possibility of investigating other features and their correlations with mobility and HIV outcomes. Additionally, the available knowledge on mobility is at the regional level in NAMPHIA, and not at the constituency level. This constraint limited the study to an inter-regional analysis rather than an intra-regional analysis.

6 Bibliography

References

[1] NAMPHIA Technical Working Group. (2018). Namibia Population-Based HIV Impact Assessment 2017 Report. https://phia.icap.columbia.edu/namphia-final-report/

[2] Palk, L. Blower, S. (2015). Mobility and circular migration in Lesotho: implications for transmission, treatment, and control of a severe HIV epidemic. J. Acquir. Immune Defic. Syndr. 68, 604–608.

[3] Elia J Mmbaga, Germana H Leyna, et al. The role of in-migrants in the increasing rural HIV-1 epidemic: results from a village population survey in the Kilimanjaro region of Tanzania. Int J Infect Dis. 2008 Sep;12(5):519-25. doi: 10.1016/j.ijid.2008.02.007. Epub 2008 May 8. PMID: 18468468.

[4] Nunn AJ, Wagner HU, et at. Migration and HIV-1 seroprevalence in a rural Ugandan population. AIDS. 1995 May;9(5):503-6. PMID: 7639976.

[5] Namibia Statistics Agency. (2015). Namibia 2011 Census Migration Report.

[6] Namibia Statistics Agency. (2019). The Namibia Labour Force Survey 2018 Report. https://www.ilo.org/surveyLib/index.php/catalog/6860.

[7] Wikipedia. Namibia. https://en.wikipedia.org/wiki/Namibia.

[8] Valdano E, Okano JT, et at. Use of mobile phone data in HIV epidemic control. Lancet HIV. 2022 Dec;9(12):e820-e821. doi: 10.1016/S2352-3018(22)00332-0. PMID: 36460021; PMCID: PMC9762893.

[9] PHIA project. https://phia.icap.columbia.edu/.

[10] Namibia Statistics Agency. (2011). Namibia 2011 Population and Housing Census.

[11] Maxmen, A. Older men and young women drive South African HIV epidemic. Nature 535, 335 (2016). https://doi.org/10.1038/nature.2016.20273.

[12] Low A, Sachathep K, et al. Migration in Namibia and its association with HIV acquisition and treatment outcomes. PLoS One. 2021 Sep 2;16(9):e0256865. doi: 10.1371/journal.pone.0256865. PMID: 34473757; PMCID: PMC8412347.

[13] Our World in Data. HIV/AIDS. Retrieved from https://ourworldindata.org/hiv-aids

[14] Aulagnier M, Janssens W, et al. (2011) Incidence of HIV in Windhoek, Namibia: Demographic and Socio-Economic Associations. PLoS ONE 6(10): e25860. https://doi.org/10.1371/journal.pone.0025860.

[15] Clark, S., Bruce, J., et al. (2006). Protecting Young Women from HIV/AIDS: The Case against Child and Adolescent Marriage. International Family Planning Perspectives, 32(2), 79–88. http://www.jstor.org/stable/4147596.

[16] Ekholuenetale M, Onuoha H,et al. (2021). Socioeconomic Inequalities in Human Immunodeficiency Virus (HIV) Sero-Prevalence among Women in Namibia: Further Analysis of Population-Based Data. International Journal of Environmental Research and Public Health. 2021; 18(17):9397. https://doi.org/10.3390/ijerph18179397.

[17] Maulsby CH, Ratnayake A, et al. (2020). A Scoping Review of Employment and HIV. AIDS Behav. 2020 Oct;24(10):2942-2955. doi: 10.1007/s10461-020-02845-x. PMID: 32246357; PMCID: PMC7716244.

[18] Olawore O, Tobian AAR,et al. (2018). Migration and risk of HIV acquisition in Rakai, Uganda: a population-based cohort study. Lancet HIV. 2018 Apr;5(4):e181-e189. doi: 10.1016/S2352-3018(18)30009-2. Epub 2018 Feb 25. PMID: 29490875; PMCID: PMC6195205.

[19] Maulsby CH, Ratnayake A, at al. A Scoping Review of Employment and HIV. (2020). AIDS Behav. 2020 Oct;24(10):2942-2955. doi: 10.1007/s10461-020-02845-x. PMID: 32246357; PMCID: PMC7716244.