UNIVERSITY OF THE WITWATERSRAND

FACULTY OF HEALTH SCIENCES

SCHOOL OF PUBLIC HEALTH

RESEARCH REPORT

Differences between 2008 and 2017 in HIV prevalence and risk factors among young adults aged 15-24 years in South Africa.

Moses Micah Mayanja

Student Number: 2388532

26 September 2025

A research report submitted in partial fulfilment of the requirements for the degree of Master of Science in Field Biostatistics in the School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg.

DECLARATION

I Moses Micah Mayanja declare that this Research Report is my own, unaided work. It is being submitted for the Degree of Master of Science in Field Biostatistics at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University. I have been trained in plagiarism and am aware that plagiarism is misconduct.

26 September 2025

DEDICATION

This work is dedicated to

My beloved parents, Steven and Cissy Mayanja, in Uganda,

whose unwavering support has been my foundation.

To my uncle and aunt, James and Annet Oguttu,

who have been my guardians throughout my MSc journey in South Africa.

And to my love, Ssiima Namusoke,

for bringing joy and love to my life.

May this degree stand as a testament to the faithfulness of God.

ABSTRACT

Background:

South Africa has a high HIV burden among young people (15-24 years). This study examines the association between HIV status and behavioural risk factors and quantifies whether changes in HIV prevalence are due to shifts in risk factor distribution or their effects.

Methods:

Data from 8,189 young adults in the 2008 and 2017 South African National HIV Surveys were analysed, by sex. Multiple imputation addressed missing data and logistic regression assessed associations between HIV and behavioural risk factors, including intergenerational sex, transactional sex, inconsistent condom use, and number of sexual partners. Decomposition analysis examined the drivers of prevalence changes.

Results:

HIV prevalence declined slightly (8.7% to 7.9%) but was not statistically significant. Intergenerational sex and early sexual debut remained key predictors. The decline in HIV prevalence was mainly driven by changes in the effects of the risk factors.

Conclusion:

Findings highlight the need for targeted interventions addressing interlinked behaviours to reduce HIV vulnerability among young adults, with emphasis on sustaining reduction among females and addressing persistent risk factors among males.

ACKNOWLEDGEMENTS

I am deeply grateful to my Lord and Saviour, Jesus Christ, whose grace has enabled to complete this degree.

I extend my sincere appreciation to my supervisor, Professor Jonathan Levin, for his invaluable support and guidance throughout the research process. His insights and expertise have been instrumental in shaping this research.

I also wish to acknowledge the late Dr. Joseph Ouma, whose pioneering idea of using decomposition analysis in studying HIV trends provided a crucial foundation for this research.

Contents

LIST OF TABLES

1. INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

This research project aims to assess the change in HIV prevalence among young adults aged 15-24 years in South Africa between 2008 and 2017, and to determine the risk factors associated with HIV infection. Furthermore, it will discuss the application of decomposition analysis to determine the extent to which changes in HIV prevalence can be attributed to changes in the distribution of risk factors as compared to changes in the effects of the risk factors. By examining these dynamics, this research seeks to provide a thorough insight into the trends and underlying important risk factors associated with HIV infections among young adult South Africans, thereby contributing to more targeted and effective interventions.

1.2 Background

Human Immunodeficiency Virus (HIV) is a viral infection specifically attacking the CD4 cells (T cells), which are important for immune defence. Without treatment, HIV can gradually destroy the immune system, causing Acquired Immunodeficiency Syndrome (AIDS), where the immune system reaches a point where it is unable to fight off infections (1), making individuals highly susceptible to opportunistic illnesses. Ever since the initial detection of HIV in the early 1980's, millions of people have been impacted worldwide and it remains a severe health burden with new infections that keep challenging the public health system.

The fight to minimize the further spread of HIV led to a 60% decrease in the number of new infections worldwide in 2023, compared to the peak in 1995 (2). The Joint United Nations Programme on HIV and AIDS (UNAIDS) report for 2023 (3) suggests that the global decrease in HIV infections is driven by a significant reduction of 57% in new HIV infections in eastern and southern Africa. These reductions in recent infections have also been attained in various other regions of the world, including the Caribbean, western and central Africa, Asia and the Pacific, North America, and central and western Europe. However, new HIV infections continued to rise in central Asia and eastern Europe, increasing by 49% since 2010.

Within Sub-Saharan Africa, despite the notable decrease in HIV infections in recent years, disparities persist within the general population (4). HIV infections are reported to be higher in females than in males. Women aged 15 years and older represented about 59% of new infections in the region (5), and specifically adult women between 15-24 years despite constituting only 10% of the population, accounted for over 25% of the new HIV infections in 2017. This gender disparity displays the need for targeted interventions that address the vulnerabilities and risk factors faced by women. For young women aged 15-24, these vulnerabilities often include gender-based violence, economic dependence on male partners, and lower negotiation power in relationships (5). These vulnerabilities can translate into

heightened behavioural and structural risk factors for HIV infection such as, economic hardship may increase susceptibility to transactional / intergenerational sex, while unequal gender power dynamics can reduce a young woman's ability to negotiate condom use or refuse unsafe sexual practices (6,7).

Since the 1990s, South Africa experienced a dramatic increase in HIV infections, marking the beginning of the country's severe HIV epidemic. In the early years there was a lack of effective response due to denial by the South African government, which greatly impacted the country's public health efforts (8). According to a study done in 2008 (9), it was estimated that during the period 2000-2005 over three hundred and thirty thousand premature deaths occurred due to the lack of effective response in the nation that could have been prevented with access to antiretroviral drugs (ARVs). Furthermore, the same study revealed that not initiating a mother-to-child transmission prevention program resulted in thirty-five thousand babies born with HIV. This led to a consistent increase in new HIV infections resulting in the rising prevalence levels in South Africa during the early years of the epidemic. Annual antenatal surveys revealed an increase in the national HIV prevalence from 0.7% in 1990 to 29.5% in 2004 (8), after which it later stabilised as the government began to take more significant steps to combat the epidemic. This was characterized by initiatives such as a substantial increase in access to antiretroviral therapy (ART), the expansion of programs like the prevention of mother to child transmission (PMTCT), and the implementation of several HIV prevention programs (8). These public health initiatives contributed to a gradual decline in new HIV infections in the early 2000s, with national surveys showing notable reductions among adults aged 15-49 (10).

In 2014 the country adopted the UNAIDS 90-90-90 targets, with the aim that 90% of people living with HIV know their status, 90% of those diagnosed and know their HIV status initiate and adhere to ART, and 90% of those on treatment achieve viral suppression. Mathematical models suggest that if these targets are achieved then the AIDS epidemic could be ended by 2030 (11). This led to improved testing and counselling services in addition to increased enrolment of infected individuals on ART. As a result of these initiatives, the epidemic stabilized with a notable decrease in the number of new HIV infections. The estimated Mother-to-child transmission rates dropped from 9.6% to 2.8% between 2008 and 2011 translating to reductions in mortality both among infants and among children under 5 years (12,13). There was also a decrease in AIDS-related deaths leading to a marked increase in life expectancy from 56 to 60 over the period 2009 – 2012 (13). However, there has been a continued rise in national HIV prevalence, from 10.6% in 2008, to 12.2% in 2012, and further to 14% in 2017 (14,15). Among specific age groups, the HIV prevalence for adults aged 15-49 has also continued to rise from 15.6% in 2002 to 16.9% in 2008, and further to 20.6% in 2017. This could be attributed to an improvement in HIV treatment; therefore, people infected with HIV are now surviving much longer. Narrowing down to adults aged 24-49, HIV prevalence was

highest in 2017 peaking at 26.4%. In contrast, HIV prevalence estimates among young adults aged 15-24 showed a decline, decreasing from 8.7% in 2008 to 7.9% in 2017 (14,15). This decline among 15-24 year olds may reflect the positive effects of intensified prevention efforts targeting youth, such as expanded HIV awareness campaigns, improved access to voluntary medical male circumcision, and school-based HIV education (16).

Behavioural characteristics are potential risk factors, reducing these risky behaviours may lead to a reduction in incident HIV infections. There are associations between HIV prevalence/acquisition and some behavioural risk factors, such as failure to use condom consistently and correctly, having multiple sex partners, and transactional sex, which have been studied and documented in the literature (17,18). For the 15-24 age group specifically, structural factors such as low school attendance / school dropout and unemployment may also further increase the risk. Young women Who attended more school days and stay in school have a lower risk of incident HIV and HSV-2 infection (19).

These findings have provided information to support resource allocation, especially in resource limited countries. However, the contrast between declining prevalence in 15-24 year olds and rising prevalence in older adults highlights the need for targeted youth-focused interventions to sustain and build on prevention gains. Information from this group is essential for shaping future policy, tailoring interventions, and maintaining downward trends in new infections. However, there is a need to carry out research to determine the extent to which the change in HIV prevalence/infections over time can be explained by the change in risky behaviour. Therefore, the study aims to investigate the contribution of changes in behavioural risk factors to changes in HIV prevalence in South Africa between 2008 and 2017.

1.3 Literature Review

HIV prevalence in South Africa is still a critical public health issue. According to a press release by the Human Sciences Research Council (HSRC) in 2023 (20), 7.8 million people are living with HIV in South Africa. Given that globally 39.9 million people were living with HIV in 2023 (2), therefore approximately 20% of all individuals living with HIV worldwide are in South Africa, underscoring the country's significant burden in the global HIV epidemic.

Understanding the factors driving the HIV epidemic in South Africa is crucial for effective response. Many factors have been identified and studied in literature. Studies indicate that there is a link between socioeconomic status and HIV outcome (21). According to the South African National HIV Prevalence, Incidence and Behaviour Survey (SABSSM) of 2012, households with a lower socioeconomic status have significantly ($p<0.05$) higher HIV prevalence compared to households of a better socioeconomic status. These households are reported to have a shortage of basic social services like water and sanitation. Generally, individuals living in poor communities are disadvantaged by a lack of access to information on

the prevention of HIV and health services such as testing for HIV infection, resulting in high HIV prevalence levels compared to upper socioeconomic communities (21).

Studies also highlight sexual high-risk behaviour factors that significantly contribute to the high rates of HIV infection. Age-disparate sexual relationships refer to a five or more-age gap between sexual partners. Systematic reviews and meta-analysis of published studies up to January 2020 in sub Saharan Africa indicated that such relationships are a high-risk factor for HIV infection (22). The 2017 SABSSM study (15) revealed that in South Africa around 35.8% of adolescent females had sexual partners who were at least five years older than themselves, as compared to only 1.5% of males. This is therefore a key determinant of high HIV incidence rates in young women mainly because of the increased possibility of older partners being HIV-positive, as well as participation in other risky behaviours such as inconsistent condom use (23). According to the World Health Organisation condoms are very effective in the prevention of sexually transmitted infections, including HIV. A study done in heterosexual couples in the United States indicated that consistent condom use is protective against HIV infection, resulting in 80% reduction in HIV infection (24). Across age groups in South Africa, consistent condom use is highest among individuals aged 15-24 years at 44.8%, then followed by individuals aged 25-49 at 26.8% and lowest at 11.3% in individuals aged 50 and older. This corresponds to HIV prevalence observed across these age groups 15-24, 25-49, 50 and older as 7.9%, 26.4% and 12.5% respectively in 2017 (15). According to researchers like Morris and Kretzschmar (25) having more than one sexual partnership significantly increased the transmission of HIV among sexually active individuals and therefore contributed to the growth rate of the epidemic in its early phase (25). Multiple sexual partners are one of the sexual risk factors that place young people at risk of HIV infection as compared to other age groups. The SABSSM studies over the years have indicated that young adults aged 15-24 have higher proportions of individuals who have reported two and more partners as compared to other age groups. In 2017, 15-24 years had 17.4% reporting to having had multiple sexual partners with 25-49 and 50 and older years reporting 10.2% and 3.6% respectively (15). Studies reveal that individuals with two or more concurrent sexual partners are very likely to participate in high-risk sexual activities like age-discrepant sexual activities, and inconsistent condom use. A study carried out among university students in South Africa revealed that students who have multiple partners are more likely to report inconsistent condom use (26). Studies indicate that early sexual debut may increase the risk of HIV infection among young people especially females (14,27). A 2009 study done in South Africa (28) found that the age at sexual debut was influenced by forced sex at first experience. Young women who experienced forced sex with their first partners were twice as likely to have had sex before the age of 15, compared to those who had not been forced. Early sexual debut and condom use at first sex were

influenced by whether the respondent's first partner had forced them, thereby increasing the risk of HIV infection.

Other than sexual high-risk factors, non-sexual risk factors like substance abuse are also linked to increased HIV infections and have been investigated and reported in the literature. According to the Center for Disease Control and Prevention (CDC) sharing needles and syringes during injection drug use puts people at high-risk of HIV infection and transmission. A study done in China where heroin is a popular drug, revealed that 42% of HIV infections in 2004 were estimated to be due to injection drug use (29). According to the UNAIDS (11), among 74 countries reporting relevant HIV prevalence statistics in 2014, people who inject drugs were 28 times more likely to have HIV than the general adult population. Reports from organisations like the National Institute on Drug Abuse (NIDA) reveal that substance abuse leads to impaired judgement, bringing about unsafe sexual practices that facilitate the spread of HIV. Research also shows that substance abuse, including alcohol abuse, may increase the risk of sexual violence among men, particularly those with a predisposition toward sexual aggression (30). This behaviour is highly correlated with increased HIV rates among women (31), who are usually victims of such violence leading to unsafe sex.

Furthermore, studies have also shown that clinical factors such as the presence of other sexually transmitted infections (STIs) significantly increase susceptibility to HIV infection (32). STIs such as syphilis and gonorrhoea can lead to open sores in the genital area, which provide the HI virus easy access to the body tissue or bloodstream, thereby increasing infectiousness during sexual contact (33). Studies reveal that individuals with STIs are more likely to engage in high-risk sexual behaviours, such as having multiple sexual relationships, which in turn increases their risk of HIV infection (34). Additionally, women with yeast and bacterial vaginal infections such as bacterial vaginosis, have an increased HIV transmission risk if exposed to sexual contact (35).

Beyond identifying risk factors, literature also demonstrates significant investment in HIV prevention interventions targeting young people. The Determined, Resilient, Empowered, AIDS-free, Mentored and Safe (DREAMS) program represents one of the most comprehensive interventions implemented to reduce HIV infections among adolescent girls and young women (AGYW) in sub-Saharan Africa. It is a combination of prevention program packages addressing biomedical, behavioural, and structural interventions including education and economic empowerment, and access to sexual and reproductive health services (5,36). Studies have demonstrated that DREAMS interventions were associated with favourable outcomes such as behavioural changes among young women, with participants more likely to practice safer sexual behaviours including consistent condom use and HIV testing when they accessed DREAMS intervention programs like school-based HIV prevention programs (37). Evaluation of this program among sub-Saharan countries has reported a

reduction in HIV incidence and STI symptoms (36) among adolescent girls and young women in the supported areas. Another critical factor that has significantly influenced HIV incidence, including among young people, is the expansion of antiretroviral therapy (ART) access and the implementation of treatment as prevention strategies. Evidence shows that when people living with HIV consistently use ART and achieve viral suppression, their risk of transmitting the virus to sexual partners is effectively eliminated (38). This population level protective effect benefits not only those on treatment but also HIV-negative individuals, including adolescents and young adults at risk of infection. This, therefore, contributes indirectly to lowering HIV incidence among youth. Furthermore, the introduction of same day ART initiation polices has enhanced the prevention impact by minimizing delays between HIV diagnosis and treatment commencement (39). This ensures earlier viral suppression and reducing opportunities for onward transmission.

Despite the broad research on the determinants of HIV infections and various interventions, there is still a significant gap in studies done to estimate the contribution of behaviour change to changes in HIV prevalence / acquisition over time. Several studies in South Africa focus on the correlation between risky behaviours and HIV prevalence; however, no studies employ methodologies that can isolate the contribution of behaviour changes and their evolving impact on HIV prevalence / acquisition over time.

1.4 Problem statement

Increased knowledge about factors associated with HIV infections has led to implementation of programs like DREAMS (Determined, Resilient, Empowered, AIDS-free, Mentored and Safe) and PMTCT (Prevention of Mother to Child Transmission), with the aim of reducing HIV infections. However, the HIV prevalence remains high especially in South Africa. According to the UNAIDS Global Aids Update, in 2017, approximately 800,000 people in Eastern and Southern Africa acquired HIV, and 380,000 died from an AIDS-related illness. South Africa accounted for 33% of the new infections and 29% of the AIDS-related deaths (5). The behavioural patterns of individuals can result in changes in the HIV prevalence. One way to understand such behavioural patterns (or their effects) is to describe their decomposition using methods developed by Blinder and Fairlie (40,41). No research has been conducted to quantify the contribution of behavioural changes and their evolving impact on the changes in HIV prevalence and HIV incidence, particularly in South Africa. Although decomposition analysis has not been widely used in Health Research, the method has been used in other fields such as Econometrics, where the decomposition analysis has been used to investigate race and sex wage differences, breaking down the overall change in wage differences between race and sex groups into portions due to the differences in factors and differences in effects of the factors (40,42).

1.5 Justification

Comparing HIV prevalence over time (2008 - 2017) is crucial for assessing the impact of HIV prevention and treatment programs, thereby identifying areas where intervention is needed. By tracking the trends of HIV prevalence through multiple survey rounds, researchers can investigate the effectiveness of public health interventions over the years. This period (2008-2017) is sufficiently long to detect change in HIV prevalence and captures the most recent available national HIV prevalence (2017).

HIV prevalence in the age group of adolescents and young adults aged 15 – 24 can be regarded as a proxy for HIV incidence in large-scale surveys (43). Young people have higher risky behaviours including early sexual debut and lower condom use rates therefore, a higher likelihood of recent HIV infections. Younger individuals also typically have shorter duration of infection compared to older age groups making them more likely to represent recent transmissions (44). Therefore, the trends in HIV incidence are effectively monitored by utilizing data on HIV prevalence among younger people.

Investigating changes in risky behaviours over time alongside changes in HIV prevalence is crucial for understanding HIV transmission dynamics and designing effective prevention strategies. Changes in risky behaviours, such as intergenerational sex, inconsistent condom use, transactional sex, or number of sexual partners in the past year, can directly influence HIV transmission rates within a population (24,45). For instance, an increase in risky behaviours, such as injected recreational drugs, can lead to a surge in new HIV infections (29), contributing to an overall increase in HIV prevalence. Therefore, by examining changes in these behaviours over time, we can assess the effectiveness of existing interventions, and target prevention efforts to address evolving challenges in HIV prevention and control.

Decomposition analysis will investigate the overall change in the effects of risky behaviours associated with HIV infection, breaking down the change in the proportion of HIV-infections into the contribution of the change in the proportion of individuals engaging in risky behaviour and the change in the impact of risky behaviours. This will highlight important risk factors associated with HIV infection and interventions that may promote positive behaviour towards HIV transmission can be implemented.

1.6 Research question

How have risky behaviours changed among young adults in South Africa, ages 15-24, between 2008 and 2017 survey years? Is any change in the prevalence of HIV over time due to a change in the proportion exhibiting the risk factor or due to a change in the impact of the risk factor?

1.7 Aim

To determine whether the change in the HIV infections between the surveys in 2008 and 2017 is attributable to a change in risky behaviour or a change in the impact of risky behaviour among adolescents and young adults aged 15-24 in South Africa.

1.8 Objectives

1. To assess the change in HIV prevalence among adolescents and young adults aged 15-24 years in South Africa between the surveys of 2008 and 2017, stratified by sex.
2. To determine the risk factors associated with HIV positivity in both 2008 and 2017 surveys, stratified by sex.
3. To quantify the separate contributions to the change in HIV prevalence between the 2008 and 2017 surveys into the portion attributable to change in risk factors and the portion attributable to change in effects of the risk factors, stratified by sex.

2. METHODS

2.1 Study setting and design

The South African national HIV prevalence, incidence, behaviour, and communication survey (SABSSM) is a nationally representative population-based survey of households in South Africa. This survey has been run periodically by the Human Sciences Research Council (HSRC) since 2002.

The primary study consists of a repeated national cross-sectional survey carried out in 2008 and 2017. For both surveys the main objectives are to estimate the HIV prevalence and incidence among children and adults in South Africa, and to evaluate the association between socio-demographic and behavioural characteristics and HIV infection (14,15). The surveys provide information on national progress towards HIV control in South Africa. A multi-stage stratified cluster random sampling design was implemented in both 2008 and 2017 surveys (14,15). From the sampling frame created by Statistics South Africa, 1000 small area layers were randomly selected after stratification by province and location type. In the 2008 survey, location type was stratified into 4 levels differing from the only 3 levels in 2017. In this study, small area layers were designated as the primary sampling units, and within each sampled small area layer, 15 households were randomly selected as the secondary sampling units. Several questionnaires were administered to households according to the age and role of the individuals within the household, with sexual health and behaviour as the primary focus of the questions. To test for HIV status, dried blood spot (DBS) samples were taken by finger prick or heel-prick and examined for HIV antibodies. More details about the 2008 and 2017 surveys are available in the study reports (14,15).

2.2 Study population, data collection and variables

Both the 2008 and 2017 surveys included household members of all ages across South Africa. The samples excluded individuals who lived in schools, nursing homes, hospitals, or police military barracks.

The primary data was collected in household surveys, with four questionnaires administered: (1) Household questionnaire; (2) Questionnaire for parent/guardian for children aged 0 to 11 years; (3) Questionnaire for children aged 12 to 14; (4) Questionnaire for persons aged 15 and older. Permission to use the secondary data was obtained from the HSRC website where the data are made publicly available.

Young adults aged 15-24 from both the 2008 and 2017 surveys are considered for the secondary data analysis. From the HSRC reports, the sample size of young adults aged 15-24 years in the 2008 survey was 5344, which when weighted up corresponds to a population

of approximately 9.9 million. In 2017 the sample size of young adults aged 15-24 years was 6854, which when weighted up corresponding to a population of approximately 9.6 million.

Variables

Outcome
HIV status (Negative, Positive)

Demographics
Age, Sex, Race, Province, Locality type

Exposures
Intergenerational sex, inconsistent condom use, transactional sex, age at sexual debut, number of sexual partners in the past year, recreational drugs use, and excessive alcohol consumption.

2.3 Data management and analysis

Both 2008 and 2017 datasets are stored and received in excel format and were imported into StataSE 17. Exploratory data analysis was used to check for internal consistency and completeness of the data. This involved investigating the extent of missing data, cleaning, and recoding variables. To address missing data in the exposure variables, multiple imputation by chained equations (MICE) was applied using logistic regression for binary variables and Poisson regression for the count variables. The demographic variables Age, Sex, Race, and Province were included as auxiliary variables to improve imputation accuracy by capturing relationships between observed and missing data. Ten imputed datasets were generated to ensure sufficient variability between imputations to produce reliable parameter estimates.

Objective 1

The datasets from 2008 and 2017 were pooled and set up for survey analysis using the “svy” commands, with sampling weights applied and clustering based on enumeration areas (EAs) in the 2008 survey and small area layers in the 2017 survey. To distinguish between the two datasets in the pooled data, a variable representing the year was added, and the cluster numbers for 2008 were modified by adding 10,000 to each. A descriptive table presents the overall weighted HIV prevalence as well as the prevalence for selected socio-demographic variables among young adults aged 15-24, illustrating the HIV prevalence at two distinct time points (2008 and 2017), with separate analysis conducted for males and females. The difference in the overall HIV prevalence between the 2008 and 2017 surveys was statistically evaluated using the univariate logistic model for HIV outcome against the survey year, stratified by sex. The null hypothesis for this test states that there is no difference in the overall HIV proportions between the two populations for each stratum.

Objective 2

Multiple logistic regression was applied separately to each survey round using the "svy" command in Stata to investigate factors associated with HIV positivity, with analyses conducted separately for males and females. Risk factors such as intergenerational sex, inconsistent condom use, transactional sex, age at sexual debut, number of sexual partners in the past year, and recreational drugs were investigated. Separate logistic regression models were used for each survey round and each sex stratum to investigate these factors. One model considered these risk factors directly as predictor variables while the second model considered the investigation of a risk score index variable generated based on these known risk factors. The derivation of the risk score was based on factors directly associated with HIV prevalence in both the 2008 and 2017 surveys. Each level of these risk factors was assigned to a score reflecting the degree of risk. The total risk score for per person was the derived by summing the scores from all n risk factors considered, using the formula:

Total Risk Score = $(Factor_1\ score) + (Factor_2\ score) + \cdots + (Factor_n\ score)$

This cumulative score provided a quantitative measure of an individual's overall risk, with higher scores indicating greater risk which was then used as a predictor in the regression model.

Objective 3

Blinder-Oaxaca decomposition analysis is a statistical method to decompose the mean difference between two group outcomes, into a portion attributable to a difference in characteristics and a portion attributable to differences in coefficients (46,47). According to Blinder (1973, p.438) the average wages for each group i.e., sex or race, can be estimated using a linear regression model:

$$\bar{Y}^k = \beta_0^k + \sum_i^n \beta_i^k \bar{X}_i^k$$

where $\bar{Y}$ represents the mean income for the group k (k= A, B) and $\bar{X}_i$ the average characteristics such as education, occupation, and years of experience. From the average outcome equation above, the average difference between these two comparison groups $\bar{Y}_A - \bar{Y}_B$, can be decomposed into a portion attributable to the difference in characteristics $\sum_i \beta_i^A \bar{X}_i^A - \sum_i \beta_i^B \bar{X}_i^B$ and the amount attributable to coefficients $\beta_0^A - \beta_0^B$. This raw difference can be further explained using the twofold and threefold decomposition equations (42). The twofold decomposition divides the raw difference into two components with a non-discriminatory coefficient vector β^*such that:

$$R = \{E(X_A) - E(X_B)\}'\beta^* + \{E(X_A)'(\beta_A - \beta^*) + E((X_B)'(\beta^* - \beta_B)\}$$

where the first component is the outcome differential explained by differences in characteristics and the second component is an unexplained part. This unexplained component encompasses differences in coefficients and all possible effects resulting from differences in unseen factors (42). The threefold decomposition divides the raw difference into three components:

$$R = \{E(X_A) - E(X_B)\}'\beta_B + E\left(X_\beta\right)'(\beta_A - \beta_B) + \{E(X_A) - E(X_B)\}'(\beta_A - \beta_B)$$

where the first component is due to difference in characteristics also known as the "endowments effects", the second component is due to differences in coefficients and lastly the third component that is an interaction term accounting for simultaneous differences in characteristics and coefficients. The Blinder-Oaxaca method is widely used to identify and quantify separate contributions of group differences in measurable characteristics but was originally developed for linear models (41). When the outcome is binary, several extensions have been proposed for the decomposition of a non-linear model. One approach is the Fairlie decomposition (48), which extends of the Blinder-Oaxaca method by differences in a binary outcome using a nonlinear equation:

$$R = \left[\sum_{i=1}^{N^A} \frac{F(X_i^A \beta^B)}{N^A} - \sum_{i=1}^{N^B} \frac{F(X_i^B \beta^B)}{N^B}\right] + \left[\sum_{i=1}^{N^A} \frac{F(X_i^A \beta^W)}{N^A} - \sum_{i=1}^{N^A} \frac{F(X_i^A \beta^B)}{N^A}\right]$$

where N^k is the sample size for group k. The first term in brackets represents the portion of differences due group differences in characteristics and the second term represents the portion due to differences in the effects of the characteristics (48). The "fairlie" package in Stata implements this method.

Another approach proposed by Yun (49) is to use the linearized approximation of the log-odds in the Blinder-Oaxaca decomposition framework. The expectation of the binary outcome in a logistic regression model is given by:

$$E[Y|X] = \frac{e^{X\beta}}{1 + e^{X\beta}}$$

the difference in group means can be rewritten using a first-order Taylor series expansion around $X\beta$, leading to the decomposition equation:

$$R = \left[\frac{\delta E[Y]}{\delta X}(X^A - X^B)\right] + \left[\frac{\delta E[Y]}{\delta \beta}(\beta^A - \beta^B)\right]$$

This approach allows the Blinder-Oaxaca decomposition to be applied in a non-linear framework. The first term represents the endowment effect (differences in characteristics), and

the second term represents the coefficient effect (differences in the impact of characteristics) (49). The “oaxaca” package in Stata implements this method.

For this study, the outcome (HIV status) is binary comparing two surveys groups, i.e., 2008 and 2017 with risky behaviour characteristics such as intergenerational sex, consistent condom use, transactional sex, age at sexual debut, and the number of sexual partners in the past year as the predictor variables, with analyses conducted separately for males and females.

The “fairlie” command in Stata does not allow the direct separation into the characteristics and coefficient effects, therefore, we applied the method proposed by Yun (49) using the “oaxaca” package with the logit option. This makes it possible to decompose differences in HIV predicted log-odds between the 2008 and 2017 surveys, explicitly distinguishing between the changes in the characteristics and the changes in coefficients, stratified by sex. This method allows either a two-fold and three-fold decomposition approach; however, the two-fold decomposition was applied for its straightforward breakdown and this enables easier interpretation. In the two-fold decomposition analysis for non-linear models, the results are presented in three key components: the overall difference, the explained component, and the unexplained component. The overall difference represents the total disparity in the predicted HIV prevalence between the 2008 and 2017 surveys for each sex stratum. A positive overall difference indicates a higher HIV prevalence in the reference group (2008), while a negative overall difference indicates an increase in HIV prevalence from the 2008 to 2017. The explained component reflects the portion of the overall difference attributable to differences in the observable characteristics between the two surveys. A positive explained component reflects an increase in the 2017 survey HIV prevalence, if the respondents had the same characteristics as the 2008 survey respondents. Conversely, a negative explained component reflects a decrease in the 2017 survey HIV prevalence, after adjusting the characteristic levels to the 2008 survey (42). Therefore, a negative explained component implies that there was an increase in the proportions of the risk factors in 2017 increasing the HIV prevalence as compared to 2008. The unexplained component represents the portion of the overall differences from the effects of these risk characteristics. A positive unexplained component suggests that the effects of risk factors had a stronger impact in 2008, contributing to a higher HIV prevalence as compared to the 2017 survey, therefore a higher overall positive difference in HIV prevalence between the two years. In contrast, a negative unexplained coefficient suggests that the effects of the risk factors on HIV prevalence intensified in 2017, therefore contributing to an increase in HIV prevalence from 2008 to 2017.

Categorical predictor variables were modelled using dummy variables for the different categories in the regression equation with the base category omitted (42,50). The decomposition results depend on the choice of the omitted base category, and if the base changes, the decomposition results change (42). To address this complexity associated with categorical predictors, we used the risk score index, that is a continuous variable with a meaningful zero corresponding to the absence of risky behaviour.

To enhance the analysis in Objective 2, the logistic regression model was expanded by pooling data from both years and incorporating the year variable along with its interactions with other predictor variables as an alternative analysis, with separate models fitted for males and females. This extension allows for an evaluation of the overall change in HIV prevalence among young adults aged 15-24 at two time points (2008 and 2017) and the change in impact of risk factors by sex. The results from this extended analysis were compared with those from the decomposition analysis for each sex stratum.

When analysing data from a complex multistage survey, we need to consider both the sampling weights and the clustering in the survey; this is done by using the “svy” commands in Stata. This will be implemented for both Objective 1 and Objective 2. Yun’s adaptation of Oaxaca decomposition is not covered by the “svy” commands; therefore, we will incorporate the sampling weights using the [pweight=wt] option. Clustering will be accounted for by using the “deft” (square root of the design effect) to inflate the standard errors for the estimated effects and hence find corrected confidence intervals for each sex-stratified analysis.

3. RESULTS

3.1 Summary of study populations in 2008 and 2017.

A total sample of 8189 young adults aged 15-24 were considered from the 2008 and 2017 surveys, with 3617 participants from the 2008 survey and 4572 participants from the 2017 survey. Participants with missing HIV status were not imputed and were therefore excluded from the analysis; as a result, the final sample included 8,189 individuals rather than the total 12,198 participants.

In 2008, the sample comprised of 1631 males (45.1%) and 1986 females (54.9%), while in 2017 it comprised of 2065 males (45.2%) and 2507 females (54.8%). Summary statistics displaying frequencies and weighted proportions of the socio-demographic characteristics, individual behaviour risk factors, and the HIV status are presented by survey round and disaggregated by sex, with separate analyses conducted for males and females. While the individual variable missing data patterns are shown in Table 1, the overall number of participants with any missing data across risk behavioural variables increased from 580 in 2008 to 1559 in 2017. Among males, missing cases rose from 259 to 651, while among females, missing cases increased from 321 to 908. Overall, the proportion of participants with incomplete data increased from 16.04% in 2008 to 34.10% in 2017.

Considering males, the age distribution shifted across the surveys. Males aged 15-19 comprised 57.8% of the weighted sample in 2008 and 47.8% in 2017, while those aged 20-24 accounted for 42.2% and 52.3% respectively. Slight changes were also observed in the racial composition of males. The proportion of Black Africans increased slightly from 80.4% in 2008 to 82.6% in 2017, indicating a marginal rise in representation. The proportion of White respondents decreased from 7.7% in 2008 to 6.9% in 2017, similarly the Indian/Asian and Coloured respondents dropped from 2.9% to 1.7%, and 9.0% to 8.8% respectively. The geographical distribution of males showed substantially changes between the two surveys. The proportion of respondents residing in urban areas increased from 62.0% in 2008 to 65.5% in 2017, similarly the rural formal (farm) area proportions increased from 3.7% to 4.8%. Conversely, the proportion of those in rural informal (tribal) areas decreased from 34.3% to 29.7%. At provincial level, the representation of male respondents from the Northern Cape, KwaZulu-Natal, Gauteng and Mpumalanga increased, whereas the proportions in Eastern Cape, Free State, and Limpopo declined.

Considering females, the age distribution shifted across the surveys. Females aged 15-19 comprised of 50.2% of the weighted sample in 2008 and 47.7% in 2017, while those aged 20-24 accounted for 49.8% and 52.3% respectively. Examining the racial compositions of females, the proportion of Coloured respondents increased from 7.4% in 2008 to 8.4% in 2017

and Indian/Asian respondents increased from 1.8% in 2008 to 2.0% in 2017. The proportion of White respondents decreased from 6.0% in 2008 to 5.3% in 2017. Regarding the geographical distribution, the proportion of female respondents in urban areas increased from 56.96% in 2008 to 64.5% while the proportion living in rural areas (formal and informal) decreased. At provincial level the proportional of female respondents from the Western Cape, Northern Cape, Gauteng and Mpumalanga increased, while the proportions in Eastern Cape, Free State, KwaZulu Natal, Northwest and Limpopo decreased.

Behavioural risk factors also exhibited notable differences between survey rounds for both males and females, and for many of these risk factors, the proportions of missing data increased from 2008 to 2017. Each risk factor was categorized into levels of "not at risk" and "at risk" based on the potential for HIV exposure. The "not at risk" category included individuals for whom the behaviour was not applicable, as well as those who engaged in the behaviour in a way that did not increase their risk. Contrary to our expectations, certain risky behaviours increased from 2008 to 2017 in both males and females. The proportion of male individuals reporting early sexual debut, intergenerational sex, transactional sex, recreational drug use and excessive alcohol consumption increased from 2008 to 2017. Specifically, the proportion of males reporting early sexual debut increased from 22.4% in 2008 to 32.7% in 2017, while the proportion of intergenerational sex engagement increased marginally from 5.2% to 5.6%. Furthermore, the prevalence of transactional sex among males also increased from 0.5% to 2.2% during this period, and reported recreational drug use more than doubled, from 7.6% to 15.8%. Excessive alcohol consumption also increased, from 15.9% in 2008 to 17.8% in 2017. The proportion of female individuals reporting intergenerational sex, transactional sex, recreational drug use and excessive alcohol consumption increased from 2008 to 2017. Specifically, the proportion of females reporting intergenerational sex increased from 14.7% in 2008 to 16.3% in 2017, while the proportion of transactional sex increased from 0.2 to 1.1%. Similarly, the prevalence of recreational drug use increased from 1.6% to 4.2%, and excessive alcohol consumption from 6.7% to 11.6% over the same period. In contrast, inconsistent condom use decreased from 13.6% in 2008 to 7.8% in 2017. These findings highlight the need for interventions to combat the rise of multiple risky behaviours, while also indicating that targeted efforts to promote safer sexual practices with condom use, is yielding positive outcomes among female young adults.

Finally, HIV prevalence showed divergent patterns by sex: among males, prevalence increased from 3.6% in 2008 to 4.8% in 2017, whereas among females it declined from 13.9% to 10.9% over the same period.

Table 1.1: Descriptive Analysis of Socio-Demographic Characteristics, Behavioural Risk Factors, and HIV Status by Survey Round for Males (SABSSM 2008 & 2017)

Variable	Levels	SABSSM 2008 n (weighted %)	SABSSM 2017 n (weighted %)
Socio-demographic			
Age (Years)	15-19	941 (57.8)	1106 (47.8)
	20-24	690 (42.2)	959 (52.3)
Race	Black African	958 (80.4)	1372 (82.6)
	White	132 (7.7)	145 (6.9)
	Coloured	376 (9.0)	394 (8.8)
	Indian/Asian	165 (2.9)	154 (1.7)
Geographical type	Urban area	1189 (62.0)	1218 (65.5)
	Rural informal (tribal areas)	373 (34.3)	594 (29.7)
	Rural formal (farms)	69 (3.7)	253 (4.8)
Province	Western Cape	262 (11.2)	241 (11.24)
	Eastern Cape	227 (13.8)	218 (10.7)
	Northen Cape	127 (1.7)	204 (2.3)
	Free State	118 (6.9)	157 (5.0)
	KwaZulu-Natal	275 (17.0)	402 (19.2)
	North West	122 (6.9)	166 (6.8)
	Gauteng	231 (20.7)	282 (25.1)
	Mpumalanga	111 (7.8)	208 (9.0)
	Limpopo	158 (13.0)	187 (10.7)
BEHAVIOUR FACTORS			
Sexual debut	Not at risk	1090 (67.1)	1229 (53.5)
	At risk (<= 16 years)	358 (22.4)	618 (32.7)
	Missing	1631 (10.4)	217 (13.8)
Intergenerational Sex	No	1378 (85.3)	1717 (79.6)
	Yes	84 (5.2)	107 (5.6)
	Missing	169 (9.5)	241 (14.8)
Consistent condom use	Not at risk	1332 (83.9)	1321 (61.0)
	At risk	118 (6.2)	114 (6.3)
	Missing	181 (9.9)	630 (32.7)
Transactional sex	No	1441 (89.8)	1805 (83.6)
	Yes	6 (0.5)	36 (2.2)
	Missing	184 (9.7)	224 (14.21)
Recreational drug use	No	1346 (84.9)	1554 (72.7)
	Yes	140 (7.6)	352 (15.8)
	Missing	145 (7.6)	159 (11.6)
Excessive alcohol	No	1169 (76.4)	1545 (70.7)
	Yes	314 (15.9)	364 (17.8)
	Missing	148 (7.7)	156 (11.5)
Number of sexual partners	n	1477	1824
	Median (IQR)	0 (0 -1)	0 (0 – 1)
	Min - Max	0 - 13	0 - 17
HIV outcome	Positive status	45 (3.6)	87 (4.8)
	Negative status	1586 (96.4)	1978 (95.2)

Risk score	0	754 (45.75)	774 (32.37)
	1	249 (14.44)	170 (8.07)
	2	214 (15.03)	274 (14.84)
	3	129 (8.30)	145 (7.74)
	4	50 (2.57)	52 (3.23)
	5	7 (0.32)	7 (0.53)
	6	1 (0.03)	0 (0.00)
	Missing	227 (13.56)	643 (33.22)

Table 2.2: Descriptive Analysis of Socio-Demographic Characteristics, Behavioural Risk Factors, and HIV Status by Survey Round for Females (SABSSM 2008 & 2017)

Variable	Levels	SABSSM 2008 n (weighted %)	SABSSM 2017 n (weighted %)
Socio-demographic			
Age (Years)	15-19	987 (50.2)	1267 (47.7)
	20-24	999 (49.8)	1240 (52.3)
Race	Black African	1337 (84.8)	1705 (84.3)
	White	120 (6.0)	134 (5.3)
	Coloured	401 (7.4)	514 (8.4)
	Indian/Asian	128 (1.8)	154 (2.0)
Geographical type	Urban area	1377 (56.96)	1511 (64.5)
	Rural informal (tribal areas)	487 (37.7)	744 (31.7)
	Rural formal (farms)	122 (5.3)	252 (3.8)
Province	Western Cape	291 (9.7)	302 (11.2)
	Eastern Cape	268 (13.7)	273 (12.0)
	Northen Cape	150 (1.6)	228 (2.1)
	Free State	120 (4.8)	177 (4.5)
	KwaZulu-Natal	343 (23.9)	512 (22.4)
	North West	152 (7.6)	226 (6.4)
	Gauteng	327 (18.6)	319 (22.8)
	Mpumalanga	144 (7.3)	262 (9.4)
	Limpopo	191 (12.8)	208 (9.2)
BEHAVIOUR FACTORS			
Sexual debut	Not at risk	(1371) 69.2	1721 (63.8)
	At risk (<= 16 years)	404 (21.2)	561 (20.3)
	Missing	211 (9.5)	225 (15.9)
Intergenerational Sex	No	1463 (75.4)	1825 (66.6)
	Yes	300 (14.7)	430 (16.3)
	Missing	223 (9.9)	252 (17.0)
Consistent condom use	Not at risk	1478 (75.6)	1414 (51.8)
	At risk	268 (13.6)	206 (7.8)
	Missing	240 (10.9)	887 (40.4)
Transactional sex	No	1739 (88.3)	2241 (82.6)
	Yes	6 (0.2)	32 (1.1)
	Missing	241 (11.5)	234 (16.2)
Recreational drug use	No	1751 (89.7)	2192 (82.0)
	Yes	43 (1.6)	144 (4.2)
	Missing	192 (8.6)	171 (13.8)

Excessive alcohol	No	1592 (84.6)	1981 (74.6)
	Yes	201 (6.7)	360 (11.6)
	Missing	193 (8.7)	166 (13.8)
Number of sexual partners	n	1786	2259
	Median (IQR)	1 (0 -1)	0 (0 – 1)
	Min - Max	0 - 29	0 – 11
HIV outcome	Positive status	234 (13.9)	249 (10.9)
	Negative status	1752 (86.1)	2258 (89.1)
Risk score	0	743 (36.62)	923 (32.81)
	1	365 (19.37)	183 (7.17)
	2	373 (20.92)	253 (9.66)
	3	175 (7.58)	197 (7.53)
	4	48 (2.55)	45 (1.61)
	5	2 (0.25)	4 (0.11)
	6	1 (0.07)	0 (0.00)
	Missing	279 (12.65)	902 (41.11)

3.2 Change in HIV prevalence between 2008 and 2017.

The overall weighted HIV prevalence among males showed an increase from 3.6% in 2008 to 4.8% in 2017, an increase of 1.2%. However, this increment was not statistically significant (P=0.42). Considering age categories, among males individuals aged 15-19 the HIV prevalence increased by 2.2% between 2008 to 2017. In contrast, among male individuals aged 20-24, there was a slight decrease in HIV prevalence by 0.3% from 2008 to 2017.

By racial group, Black African males experienced an increase in HIV prevalence from 4.5% in 2008 to 5.3% in 2017, an increment of 0.8%. Other groups, such as White, Coloured and Indian/Asian all showed increases in prevalence. This increase could be explained by low baseline prevalence values in 2008.

HIV prevalence of males in all geographic settings. In urban areas the prevalence slightly increased by 0.4%, while in rural informal areas, prevalence had an increase by 1.6%, from 2.6% to 4.2%. In rural formal (farm) areas, the change was largest by just 6.9%, from 2.4% to 9.3%. Rural formal areas have the highest HIV prevalences as compared to Urban and rural informal areas. Among provinces, KwaZulu-Natal saw a substantial decrease in HIV prevalence from 7.9% in 2008 to 5.7% in 2017, a drop of 2.2%. In contrast, provinces like Eastern Cape, Free State, Limpopo and the Western Cape experienced increases in HIV prevalence, with changes of 4.2%, 1.5% ,3.3% and 2.6% respectively.

Examining females, the overall weighted HIV prevalence indicated a decrease from 13.9% in 2008 to 10.9% in 2017, a reduction of 3%. This decrease was significant (P=0.04). With regards to age categories among female individuals aged 15-19 the HIV prevalence slightly decreased by 0.9% between 2008 to 2017, while among those aged 20-24 there was a notable decrease by 5.5% from 2008 to 2017.

Considering racial groups, Black African females experienced a decrease in HIV prevalence by 3.8% from 2008 to 2017, while other racial groups all indicated a slight increase in prevalences. HIV prevalence of females in urban and rural formal areas substantially decreased by 4.3% and 5.6% respectively from 2008 to 2017. Among provinces, Gauteng and Mpumalanga saw a substantial decrease in HIV prevalence by 6.2% and 7.6% respectively from 2008 to 2017. In contrast, provinces like Eastern Cape and Free State experienced increases in HIV prevalences, with changes of 6.2% and 1.6%.

Table 3.1: Change in HIV prevalence among adolescents and young adults aged 15-24 years in South Africa between the surveys of 2008 and 2017 for males

Characteristic	HIV prevalence (weighted) %		Percentage point difference in HIV prevalence
	SABSSM 2008	SABSSM 2017	
Overall	**3.6**	**4.8**	**- 1.2 (P-value = 0.42)**
Age (Years)			
15-19	2.5	4.7	- 2.2
20-24	5.1	4.8	0.3
Race			
Black African	4.5	5.3	- 0.8
White	0	3.3	- 3.3
Coloured	0	1.8	- 1.8
Indian/Asian	0	0.2	- 0.2
Geographical type			
Urban area	4.3	4.7	- 0.4
Rural informal (tribal areas)	2.6	4.2	- 1.6
Rural formal (farms)	2.4	9.3	- 6.9
Province			
Western Cape	0.4	3.0	- 2.6
Eastern Cape	0.8	5.0	- 4.2
Northen Cape	2.0	2.3	- 0.3
Free State	1.2	2.7	- 1.5
KwaZulu-Natal	7.9	5.7	2.2
North West	2.9	3.5	- 0.6
Gauteng	6.3	4.8	1.5
Mpumalanga	5.0	8.8	- 3.8
Limpopo	0.3	3.6	- 3.3

Table 4.2: Change in HIV prevalence among adolescents and young adults aged 15-24 years in South Africa between the surveys of 2008 and 2017 for females

Characteristic	HIV prevalence (weighted) %		Percentage point difference in HIV prevalence
	SABSSM 2008	SABSSM 2017	
Overall	**13.9**	**10.9**	**3 (P-value = 0.04)**
Age (Years)			
15-19	6.7	5.8	0.9
20-24	21.1	15.6	5.5
Race			
Black African	16.1	12.3	3.8
White	0.1	2.5	- 2.4
Coloured	2.6	4.5	- 1.9
Indian/Asian	0.05	0.8	- 0.75
Geographical type			
Urban area	13.3	9.0	4.3
Rural informal (tribal areas)	14.0	14.5	- 0.5
Rural formal (farms)	19.2	13.6	5.6
Province			
Western Cape	6.0	5.8	0.2
Eastern Cape	12.6	18.8	- 6.2
Northen Cape	6.0	6.7	- 0.7
Free State	7.5	9.1	- 1.6
KwaZulu-Natal	21.1	12.1	9
North West	9.5	9.6	0.1
Gauteng	14.5	8.3	6.2
Mpumalanga	23.1	15.5	7.6
Limpopo	7.8	8.4	- 0.6

3.3 Risk factors associated with HIV positivity in 2008 adjusting for socio-demographic variables, stratified by sex.

Individual risk factor analyses

Considering the 3617 young adults aged 15-24 captured in the 2008 survey, separate logistic regression models for males and females, accounting for clustering and weights, were used to assess each risk-factor adjusting for socio-demographic variables age, , race, province and geographical type.

Among females, those with early sexual debut had approximately twice the odds of being HIV positive when compared to those with later sexual debut adjusting for socio-demographic variables (aOR 2.04; 95%CI: 1.30 - 3.19).

Similarly, females that engaged in intergenerational sex had about twice the odds of being HIV positive when compared to those who did not engage in intergenerational sex (aOR 2.04; 95%CI: 1.24 – 3.34).

Among males, none of the risk factors were significantly associated with HIV status in 2008 adjusting for socio-demographic variables.

Table 5.1: Logistic regression models for HIV status in 2008 adjusting for socio-demographic variables for males

Variable		**Odds ratio**	**95% Conf. Interval**		**P - value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<= 16yrs)	1.00	0.33	3.09	1.00
Intergenerational sex	No	**Ref**	-	-	-
	Yes	1.43	0.28	7.27	**0.44**
Consistent condom use	Not at risk	**Ref**	-	-	-
	At risk	0.88	0.22	3.44	0.85
Recreational drug use	No	**Ref**	-	-	-
	Yes	1.44	0.25	8.32	0.68
Excessive alcohol	No	**Ref**	-	-	-
	Yes	1.19	0.34	4.13	0.27
Number of sexual partners		1.05	0.70	1.57	0.82

Table 6.2: Logistic regression models for HIV status in 2008 adjusting for socio-demographic variables for females

Variable		**Odds ratio**	**95% Conf. Interval**		**P - value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<= 16yrs)	2.04	1.30	3.19	**<0.001**
Intergenerational sex	No	**Ref**	-	-	-
	Yes	2.04	1.24	3.34	0.01
Consistent condom use	Not at risk	**Ref**	-	-	-
	At risk	1.10	0.64	1.90	0.73
Transactional sex	No	**Ref**	-	-	-
	Yes	0.71	0.35	14.07	0.82
Recreational drug use	No	**Ref**	-	-	-
	Yes	1.11	0.18	6.89	0.91
Excessive alcohol	No	**Ref**	-	-	-
	Yes	1.26	0.55	2.87	0.58
Number of sexual partners		1.02	0.90	1.17	0.73

Multivariate analysis considering predictor variables among males.

The final multiple logistic regression model adjusting for the confounding effect of socio-demographic variables and intergenerational sex, indicates that young male adults with an early sexual debut in 2008 are 6% less likely to have HIV positive status than those with a later sexual debut (aOR 0.94; 95%CI: 0.30 – 2.96). The association is not statistically significant with a p-value of 0.91.

Young male adults who engaged in intergenerational sex, adjusting for sexual debut and socio-demographic variables, are 47% more likely to have HIV positive status compared to those who do not engage in intergenerational sex (aOR 1.47; 95%CI: 0.27-8.01). This association is also not statistically significant with a p-value of 0.65.

Table 7.1: Multivariate logistic regression for HIV status in 2008 considering predictors and adjusting for socio-demographic variables for males

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<=16yrs)	0.94	0.30	2.96	0.91
Intergenerational sex	No	**Ref**	-	-	-
	Yes	1.47	0.27	8.01	0.65
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.59	0.19	1.78	0.35
	Rural formal (farms)	0.72	0.14	3.73	0.70
Race	Black African	**Ref**	-	-	-
	White	1	-	-	-
	Coloured	0.03	0.004	0.29	<0.001
	Indian/Asian	1	-	-	-
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.22	0.09	15.98	0.88
	Northen Cape	4.75	0.49	45.87	0.18
	Free State	1.55	0.23	10.54	0.66
	KwaZulu-Natal	13.18	2.49	69.79	0.00
	North West	3.75	0.57	24.61	0.17
	Gauteng	7.42	1.14	48.42	0.04
	Mpumalanga	6.29	0.78	51.01	0.09
	Limpopo	0.37	0.04	3.55	0.49
Age		1.09	0.83	1.43	0.54

Multivariate analysis considering predictor variables among females.

The final multiple logistic regression model adjusting for the confounding effect of socio-demographic variables and intergenerational sex, indicates that young female adults with an early sexual debut in 2008 are 87.0% more likely to have HIV positive status than those with a later sexual debut (aOR 1.87; 95%CI: 1.17 – 3.00). The association is statistically significant with a p-value of 0.01.

Young female adults who engaged in intergenerational sex, adjusting for sexual debut and socio-demographic variables, are 85.0% more likely to have HIV positive status compared to those who do not engage in intergenerational sex (aOR 1.85; 95%CI: 1.12 – 3.07).

Table 8.2: Multivariate logistic regression for HIV status in 2008 considering predictors and adjusting for socio-demographic variables for females

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<=16yrs)	1.87	1.17	3.00	0.01
Intergenerational sex	No	**Ref**	-	-	-
	Yes	1.85	1.12	3.07	0.02
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.89	0.55	1.44	0.63
	Rural formal (farms)	1.21	0.67	2.21	0.53
Race	Black African	**Ref**	-	-	-
	White	0.01	0.00	0.05	<0.001
	Coloured	0.18	0.08	0.38	<0.001
	Indian/Asian	0.00	0.00	0.01	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.74	0.62	4.90	0.29
	Northen Cape	1.19	0.39	3.64	0.77
	Free State	0.73	0.24	2.19	0.57
	KwaZulu-Natal	2.70	1.04	7.00	0.04
	North West	1.01	0.33	3.07	0.99
	Gauteng	1.99	0.77	5.15	0.16
	Mpumalanga	3.33	1.28	8.69	0.01
	Limpopo	0.80	0.25	2.53	0.70
Age		1.31	1.21	1.43	<0.001

Logistic regression considering risk score index.

A risk score was derived using the following risk factors; sexual debut, intergenerational sex, consistent condom use, and number of sexual partners. Each risk factor was grouped into at-risk and not-at-risk groups and assigned scores accordingly. Specifically, intergenerational sex was scored as 1 for "Yes" and 0 for "No." Consistent condom use was scored as 1 for individuals classified as at risk (i.e., those who did not consistently use condoms) and 0 for those not at risk. Sexual debut was scored as 1 for individuals whose first sexual encounter occurred at age 16 or younger and 0 for those who initiated sexual activity at 17 years or older. The number of sexual partners was categorized into four groups, with scores assigned as follows: none = 0, one partner = 1, two to five partners = 2, and six or more partners = 3.

The total risk score for per individual was calculated by summing the scores from all four variables, as described in by the formula in Chapter 2.

From the logistic regression models adjusting for socio-demographic variables, the risk score significantly increases the risk HIV positive status among females by 34% per unit increase (aOR 1.34; 95% CI: 1.09 – 1.63). In contrast, the association was not statistically significant among males (aOR 1.08; 95% CI: 0.67 – 1.73).

Table 9.1: Logistic regression for HIV status in 2008 considering risk score index and adjusting for socio-demographic variables for males

Variables		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk Score index		1.08	0.67	1.73	0.76
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.60	0.19	1.84	0.38
	Rural formal (farms)	0.75	0.14	3.92	0.73
Race	Black African	**Ref**	-	-	-
	White	-	-	-	-
	Coloured	0.36	0.00	0.30	<0.001
	Indian/Asian	-	-	-	-
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.25	0.96	16.23	0.87
	Northen Cape	4.99	0.51	48.54	0.17
	Free State	1.57	0.23	10.74	0.65
	KwaZulu-Natal	13.65	2.59	71.94	<0.001
	North West	3.93	0.61	25.36	0.15
	Gauteng	7.89	1.19	52.19	0.03
	Mpumalanga	6.61	0.86	50.64	0.07
	Limpopo	0.38	0.40	3.65	0.40
Age		1.09	0.84	1.40	0.53

Table 10.2: Logistic regression for HIV status in 2008 considering risk score index and adjusting for socio-demographic variables for females

Variables		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk Score index		1.34	1.09	1.63	0.01
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.92	0.58	1.48	0.74
	Rural formal (farms)	1.26	0.70	2.28	0.44
Race	Black African	**Ref**	-	-	-
	White	0.01	0.00	0.05	<0.001
	Coloured	0.17	0.78	0.38	<0.001
	Indian/Asian	0.00	0.00	0.01	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.65	0.58	4.66	0.34
	Northen Cape	1.08	0.35	3.35	0.90
	Free State	0.65	0.21	2.03	0.46
	KwaZulu-Natal	2.58	1.00	6.66	0.05
	North West	0.93	0.31	2.84	0.90
	Gauteng	1.84	0.70	4.79	0.21
	Mpumalanga	3.06	1.17	8.04	0.02
	Limpopo	0.71	0.22	2.28	0.57
Age		1.28	1.18	1.40	<0.001

3.4 Risk factors associated with HIV positivity in 2017 adjusting for socio-demographic variables

Individual risk factor analyses

Considering the 4572 young adults aged 15-24 captured in the 2017 survey, separate logistic regression models for males and females were employed to assess each risk-factor adjusting for socio-demographic variables age, sex, race, province and geographical type.

Among females, those with an early sexual debut, adjusted for socio-demographic variables, had more than twice the odds of being HIV positive when compared to those with a later sexual debut (aOR 2.35; 95%CI: 1.37-3.69). This association is statistically significant, with a p-value of <0.001.

Similarly females who engaged in intergenerational sex, adjusted for socio-demographic variables, are 68% more likely to have HIV positive status than those who do not engage in intergenerational sex (aOR 1.68; 95%CI: 1.00 - 2.81 0.96 - 2.42). This association is statistically significant, with a p-value of 0.05.

Among males, none of the risk factors were significantly associated with HIV status in 2017 adjusting for socio-demographic variables.

Table 11.1: Logistic regression for HIV status in 2017 adjusting for socio-demographic variables for males

Variable		**Odds ratio**	**95% Conf. Interval**		**P - value**
Sexual Debut	Not at Risk	**Ref**	-	-	-
	Risk (<= 16yrs)	0.68	0.36	1.29	0.23
Intergenerational sex	No	**Ref**	-	-	-
	Yes	0.89	0.25	3.17	0.86
Consistent condom use	Not at risk	**Ref**	-	-	-
	At risk	1.47	0.47	4.59	0.50
Transactional sex	No	**Ref**	-	-	-
	Yes	0.54	0.07	4.44	0.57
Recreational drug use	No	**Ref**	-	-	-
	Yes	0.50	0.18	1.44	0.20
Excessive alcohol	No	**Ref**	-	-	-
	Yes	1.20	0.51	2.79	0.68
Number of sexual partners		0.79	0.50	1.23	0.28

Table 12.2: Logistic regression for HIV status in 2017 adjusting for socio-demographic variables for females

Variable	**Odds ratio**	**95% Conf. Interval**	**P - value**

Sexual Debut	Not at Risk	**Ref**	-	-	-
	Risk (<= 16yrs)	2.35	1.37	3.69	<0.001
Intergenerational sex	No	**Ref**	-	-	-
	Yes	1.68	1.00	2.81	0.05
Consistent condom use	Not at risk	**Ref**	-	-	-
	At risk	1.25	0.62	2.49	0.52
Transactional sex	No	**Ref**	-	-	-
	Yes	1.99	0.48	8.26	0.34
Recreational drug use	No	**Ref**	-	-	-
	Yes	1.61	0.64	4.03	0.31
Excessive alcohol	No	**Ref**	-	-	-
	Yes	0.91	0.49	1.66	0.75
Number of sexual partners		1.38	0.96	1.97	0.08

Multivariate analysis considering predictor variables among males.

The final multiple logistic regression model adjusting for the confounding effect of socio-demographic variables, indicates that young male adults with an early sexual debut in 2017 are 32% less likely to have HIV positive status times than those with a later sexual debut (aOR 0.68; 95%CI: 0.36 - 1.28). The association is not statistically significant with a p-value of 0.23. Young male adults who engaged in intergenerational sex, adjusting for sexual debut and socio-demographic variables, are 4% less likely to have HIV positive status compared to those who did not engage in intergenerational sex (aOR 0.96; 95% CI: 0.27 – 3.42)

Table 13.1: Logistic regression for HIV status in 2017 considering predictors and adjusting for socio-demographic variables for males

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<=16yrs)	0.68	0.36	1.28	0.23

Intergenerational sex	No	**Ref**	-	-	-
	Yes	0.96	0.27	3.42	0.96
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.62	0.33	1.18	0.15
	Rural formal (farms)	1.75	0.90	3.39	0.10
Race	Black African	**Ref**	-	-	-
	White	0.47	0.10	2.14	0.33
	Coloured	0.33	0.10	1.16	0.08
	Indian/Asian	0.03	0.00	0.23	0.00
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.31	0.24	7.20	0.75
	Northen Cape	0.72	0.14	3.74	0.70
	Free State	0.62	0.12	3.32	0.58
	KwaZulu-Natal	1.49	0.32	6.99	0.61
	North West	0.79	0.14	4.39	0.79
	Gauteng	1.14	0.23	5.80	0.87
	Mpumalanga	2.53	0.50	12.80	0.26
	Limpopo	0.98	0.18	5.37	0.99
Age		1.05	0.43	0.93	1.19

Multivariate analysis considering predictor variables among females.

The final multiple logistic regression model adjusting for the confounding effect of socio-demographic variables and intergenerational sex, indicates that young female adults with an early sexual debut in 2017 twice more likely to have HIV positive status than those with a later

sexual debut (aOR 2.10; 95%CI: 1.26 – 3.50). The association is statistically significant with a p-value of 0.01.

Young female adults who engaged in intergenerational sex, adjusting for sexual debut and socio-demographic variables, are 42.0% more likely to have HIV positive status compared to those who do not engage in intergenerational sex (aOR 1.42; 95%CI: 0.83 – 2.40). This association is not statistically significant (p-value = 0.19)

Table 14.2: Logistic regression for HIV status in 2017 considering predictors and adjusting for socio-demographic variables for females

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<=16yrs)	2.10	1.26	3.50	0.01
Intergenerational sex	No	**Ref**	-	-	-
	Yes	1.42	0.83	2.40	0.19
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	1.39	0.89	2.18	0.14
	Rural formal (farms)	1.33	0.76	2.31	0.32
Race	Black African	**Ref**	-	-	-
	White	0.29	0.06	1.35	0.11
	Coloured	0.46	0.23	0.93	0.03
	Indian/Asian	0.08	0.02	0.36	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	2.39	1.09	5.25	0.03
	Northen Cape	1.03	0.33	3.27	0.96
	Free State	1.66	0.71	3.89	0.25
	KwaZulu-Natal	1.71	0.80	3.63	0.16
	North West	1.08	0.48	2.43	0.86
	Gauteng	1.26	0.58	2.71	0.56
	Mpumalanga	2.51	1.20	5.25	0.02
	Limpopo	0.93	0.40	2.14	0.86
Age		1.25	1.15	1.35	<0.001

Logistic regression considering risk score index.

A risk score was calculated in the same way as in 2008 and using the same risk factors; sexual debut, intergenerational sex, consistent condom use, and number of sexual partners.

From the logistic regression models adjusting for socio-demographic variables, the risk score increases the risk of HIV positive status by 34% per unit increase in females (aOR 1.34; 95%CI: 1.07 – 1.67). In contrast the association was risk score decreases the risk of HIV positive status by 15.0% per unit increase in males (aOR 0.85; 95% CI: 0.63-1.14)

Table 15.1: Logistic regression for HIV status in 2017 considering risk score index and adjusting for socio-demographic variables for males

Variables		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk Score		0.85	0.63	1.14	0.27
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.62	0.33	1.16	0.14
	Rural formal (farms)	1.72	0.89	3.32	0.11
Race	Black African	**Ref**	-	-	-
	White	0.48	0.11	2.17	0.34
	Coloured	0.33	0.10	1.16	0.09
	Indian/Asian	0.03	0.00	0.23	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.32	0.23	7.46	0.75
	Northen Cape	0.72	0.14	3.79	0.70
	Free State	0.62	0.11	3.45	0.58
	KwaZulu-Natal	1.48	0.31	7.16	0.62
	North West	0.81	0.14	4.60	0.81
	Gauteng	1.13	0.21	5.95	0.89
	Mpumalanga	2.56	0.40	13.32	0.27
	Limpopo	1.02	0.18	5.77	0.98
Age		1.07	0.94	1.22	0.28

Table 16.2: Logistic regression for HIV status in 2017 considering risk score index and adjusting for socio-demographic variables for females

Variables		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk Score		1.34	1.07	1.67	0.01
	Urban	**Ref**	-	-	-

Geographical type	Rural informal (tribal areas)	1.45	0.93	2.26	0.10
	Rural formal (farms)	1.40	0.80	2.45	0.24
Race	Black African	**Ref**	-	-	-
	White	0.29	0.06	1.34	0.11
	Coloured	0.47	0.24	0.94	0.03
	Indian/Asian	0.09	0.02	0.37	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	2.39	1.10	5.21	0.03
	Northen Cape	1.04	0.33	3.26	0.95
	Free State	1.64	0.70	3.86	0.26
	KwaZulu-Natal	1.66	0.79	3.49	0.18
	North West	1.03	0.46	2.32	0.95
	Gauteng	1.19	0.56	2.54	0.65
	Mpumalanga	2.36	1.13	4.94	0.02
	Limpopo	0.99	0.38	2.02	0.75
Age		1.21	1.11	1.32	<0.001

3.5 Risk factors associated with HIV positivity for pooled survey data.

Decomposition analysis of risk predictors among males.

The overall difference in male HIV prevalence between 2008 and 2017 is -1.15%, however; this difference is not statistically significant (P-value = 0.3). The difference in the P-value (0.3 vs 0.42 in Section 3.2) is because this analysis adjusts for other factors (sexual debut and intergenerational sex), while the analysis in Section 3.2 does not include any adjustments.

However, when considering the explained component in the decomposition analysis, specifically accounting for the significant predictors sexual debut and intergenerational sex in the decomposition analysis of males, we would expect the 2008 survey to have higher HIV prevalence compared to the 2017 survey. This suggests that if the 2017 male survey group had the same distribution of sexual debut and intergenerational sex as those in the 2008 survey, their HIV prevalence would have been higher (P=0.70).

The unexplained component represents the contribution of the differences in the effects (coefficients) of the risky behaviours on HIV prevalence between the survey years. This suggests that over time, the impact of sexual debut and intergenerational sex caused higher HIV prevalence in 2017, contributing to the increase in the HIV prevalence over time. However, this effect is not statistically significant (P=0.30).

Therefore, the increase in males HIV prevalence between 2008 and 2017 is primarily driven by the changes in the impact of risk factors on HIV over time.

Table 17.1: Decomposition analysis of HIV status using risk predictors for males

Decomposition analysis		**Coefficient**	**95% Conf. Interval**		**P - value**
Overall	Difference	- 0.0115	- 0.034	0.011	0.30
	Explained	0.0005	- 0.002	0.003	0.70
	Unexplained	- 0.0120	- 0.034	0.010	0.30
Explained	Sexual Debut	0.0006	- 0.002	0.003	0.60
	Intergenerational sex	- 0.0001	- 0.001	0.001	0.80
Unexplained	Sexual Debut	0.0033	- 0.012	0.019	0.60
	Intergenerational sex	0.001	- 0.003	0.005	0.80

Decomposition analysis of risk predictors among females.

The overall difference in female HIV prevalence between 2008 and 2017 is 2.96%, and this difference is statistically significant (P-value <0.001).

Considering the explained component in the decomposition analysis, specifically accounting for the significant predictors sexual debut and intergenerational sex in the decomposition analysis of females, we would expect the 2008 survey to have lower HIV prevalence compared

to the 2017 survey. This suggests that if the 2017 survey group had the same distribution of sexual debut and intergenerational sex as those in the 2008 survey, their HIV prevalence would have been lower (P=0.20).

The unexplained component suggests that over time, the impact of sexual debut and intergenerational sex caused lower HIV prevalence in 2017, positively contributing to the reduction in the HIV prevalence over time. However, this effect is statistically significant (P <0.001).

Therefore, the decline in HIV prevalence between 2008 and 2017 is primarily driven by the changes in the impact of risk factors on HIV over time.

Table 18.2: Decomposition analysis of HIV status using risk predictors for females

Decomposition analysis		**Coefficient**	**95% Conf. Interval**		**P - value**
Overall	Difference	0.0296	0.001	0.059	<0.001
	Explained	- 0.0046	- 0.011	0.002	0.20
	Unexplained	0.0342	0.005	0.063	<0.001
Explained	Sexual Debut	- 0.0004	- 0.005	0.004	0.80
	Intergenerational sex	- 0.0041	- 0.009	0.001	0.10
Unexplained	Sexual Debut	- 0.0052	- 0.020	0.010	0.50
	Intergenerational sex	0.0050	- 0.009	0.019	0.50

Decomposition analysis of risk score index among males.

The HIV prevalence among males increased by 1.15% from 2008 to 2017. An alternative decomposition analysis using the risk score indicates that the explained component has a non-significant negative coefficient (P=1.00), suggesting that, based on the observable characteristics, respondents in the 2008 survey were expected to have a lower HIV prevalence compared to those in the 2017 survey. In other words, if the individuals surveyed in 2017 had

the same distribution of risk characteristics as those in 2008, the prevalence for HIV in 2017 group would have been significantly lower. Similarly, the unexplained component has a non-significant (P=0.30) negative coefficient. This suggests that the differences in the impact of the risky behaviours negatively contributed to the reduction in HIV prevalence between 2008 and 2017.

Therefore, the increase in HIV prevalence between 2008 and 2017 is driven by both the changes in the risk factors and the impact of risk factors on HIV over time. This highlights the need to address the increasing shifts in population risk factors and the effects of these risk factors.

Table 19.1: Decomposition analysis of HIV status using risk score index for males

Decomposition analysis		**Coefficient**	**95% Conf. Interval**		**P - Value**
Overall	Difference	- 0.0115	- 0.034	0.011	0.30
	Explained	- 0.0001	- 0.002	0.002	1.00
	Unexplained	- 0.0115	- 0.033	0.010	0.30
Explained	Risk score	- 0.0001	- 0.002	0.002	1.00
Unexplained	Risk score	- 0.010	- 0.007	0.002	0.20

Decomposition Analysis of risk score index among females.

The HIV prevalence among females decreased by 2.96% from 2008 to 2017. An alternative decomposition analysis using the risk score indicates that the explained component has a non-significant negative coefficient (P=0.80), suggesting that, based on the observable characteristics, respondents in the 2008 survey were expected to have a lower HIV prevalence compared to those in the 2017 survey. In other words, if the individuals surveyed in 2017 had the same distribution of risk characteristics as those in 2008, the prevalence for HIV in 2017 group would have been significantly lower. In contrast, the unexplained component has a significant (P <0.001) positive coefficient. This suggests that the differences in the impact of the risky behaviours positively contributed to the reduction in HIV prevalence between 2008 and 2017.

Therefore, the decrease in HIV prevalence between 2008 and 2017 is driven by the changes in the impact of risk factors on HIV over time. This highlights the need to address the increasing shifts in population risk factors which have hindered an even more substantial reduction in HIV prevalence.

Table 20.2: Decomposition analysis of HIV status using risk score index for females

Decomposition analysis		Coefficient	95% Conf. Interval		P - Value
Overall	Difference	0.0296	0.001	0.059	<0.001
	Explained	- 0.0010	- 0.008	0.006	0.80
	Unexplained	0.0306	0.002	0.060	<0.001
Explained	Risk score	- 0.0010	- 0.008	0.006	0.80
Unexplained	Risk score	0.0013	- 0.024	0.027	0.90

Multiple logistic regression analysis for pooled survey data considering predictors.

Considering the 8189 young adults aged 15-24 captured in the pooled survey data from 2008 and 2017, separate gender-stratified multiple logistic regression models including the survey year variable were employed to assess risk-factors. Two separate models were fitted for males (n=3696) and females (n=4493).

Among young adult males, the final multiple logistic model adjusting for confounding effect of socio-demographic characteristics and intergenerational sex, indicated that early sexual debut was not significantly associated with HIV infection (aOR 0.70; 95%CI: 0.38 – 1.28), Engagement in intergenerational sex also showed no significant association (aOR 1.01; 95%CI: 0.43 – 2.38). A difference was observed regarding HIV prevalence between survey years. With adjusting for socio-demographic characteristics and sexual risk factors, young male adults in 2017 were 21% more likely to be HIV-positive compared to those in 2008 (aOR 1.12; 95%CI: 0.66 – 2.21)

The pattern in females differed substantially. Early sexual debut was significantly associated with increased HIV risk, with females nearly twice as likely to be HIV positive compared to those with later sexual debut (aOR 1.94; 95%CI: 1.44 – 2.61). Engagement in intergenerational sex was also significantly associated with HIV infection (aOR 1.60; 95%CI: 1.21 – 2.29). Notably, females in 2017 were 32% less likely to be HIV positive compared to those in 2008 (aOR0.68; 95%CI: 0.51 – 0.90).

Overall, these findings indicate that HIV risk factors are more pronounced and statistically significant among females compared to males, highlighting important gender differences in vulnerability to HIV infection.

Table 21.1: Multivariate logistic regression for HIV status in both 2008 and 2017, considering predictors and adjusting for socio-demographic variables for males.

Variable		Odds ratio	95% Conf. Interval		P-value
Sexual Debut	Not at risk	Ref	-	-	-
	Risk (<=16yrs)	0.70	0.38	1.28	0.20
Intergenerational	No	Ref	-	-	-

sex	Yes	1.01	0.43	2.38	0.97
Survey year	2008	Ref	-	-	-
	2017	1.21	0.66	2.21	0.54
Geographical type	Urban	Ref	-	-	-
	Rural informal (tribal areas)	0.59	0.31	1.12	0.11
	Rural formal (farms)	1.24	0.66	2.33	0.51
Race	Black African	Ref	-	-	-
	White	0.24	0.05	1.10	0.07
	Coloured	0.23	0.08	0.65	0.01
	Indian/Asian	0.01	0.00	0.06	<0.001
Province	Western Cape	Ref	-	-	-
	Eastern Cape	1.07	0.24	4.80	0.93
	Northen Cape	1.15	0.27	4.93	0.85
	Free State	0.67	0.16	2.76	0.58
	KwaZulu-Natal	2.99	0.79	11.37	0.11
	North West	1.15	0.27	4.83	0.85
	Gauteng	1.92	0.45	8.14	0.38
	Mpumalanga	2.94	0.69	12.59	0.15
	Limpopo	0.73	0.16	3.37	0.69
Age		1.08	0.95	1.23	0.26

Table 22.2: Multivariate logistic regression for HIV status in both 2008 and 2017, considering predictors and adjusting for socio-demographic variables for females

Variable		**Odds ratio**		**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	Ref		-	-	-
	Risk (<=16yrs)	1.94		1.44	2.61	<0.001
Intergenerational sex	No	Ref		-	-	-
	Yes	1.60		1.12	2.29	0.01
Survey year	2008	Ref		-	-	-

	2017	0.68		1.22	1.36	<0.001
Geographical type	Urban	Ref		-	-	-
	Rural informal (tribal areas)	1.09		0.78	1.52	0.62
	Rural formal (farms)	1.26		0.83	1.91	0.27
Race	Black African	Ref		-	-	-
	White	0.09		0.02	0.41	<0.001
	Coloured	0.30		0.18	0.51	<0.001
	Indian/Asian	0.02		0.01	0.09	<0.001
Province	Western Cape	Ref		-	-	-
	Eastern Cape	2.05		1.07	3.93	0.03
	Northen Cape	1.14		0.49	2.52	0.80
	Free State	1.06		0.52	2.13	0.87
	KwaZulu-Natal	2.22		1.22	4.07	0.01
	North West	1.02		0.51	2.03	0.96
	Gauteng	1.56		0.85	2.86	0.15
	Mpumalanga	2.93		1.61	5.35	<0.001
	Limpopo	0.83		0.40	1.73	0.62
Age		1.29		1.22	1.36	<0.001

Multiple logistic regression analysis for pooled survey data considering the risk score.

The multiple logistic regression analysis of the pooled survey data among males indicates that higher risk score was associated with lower likelihood of being HIV-positive, this however, was not statistically significant (aOR 0.90; 95%CI: 0.73 – 1.12). Males in 2017 had an 18% increased likelihood of being HIV-positive compared to those in 2008, although this association did not reach statistical significance (aOR 1.18; 95%CI: 0.64 – 2.17). This difference indicates an increased in HIV prevalence over the years in males.

By contrast, analysis in females indicates that a higher risk score is significantly associated with an increased likelihood of being HIV-positive, with adjusted odds ratio (aOR) of 1.33

(95%CI: 1.18 – 1.50). Additionally, females in 2017 were significantly less likely (30%) to be HIV positive compared to those in 2008 (aOR 0.70; 95%CI 0.53 – 0.92). This difference indicates a possible decline in HIV prevalence over time, potentially due to improved prevention strategies, greater awareness, and increased access to treatment and testing services.

Table 23.1: Multivariate logistic regression for HIV status in both 2008 and 2017, considering composite risk score and adjusting for socio-demographic variables for males

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk score		0.90	0.73	1.12	0.34
Survey year	2008	Ref	-	-	-
	2017	1.18	0.64	2.17	0.59
Geographical type	Urban	Ref	-	-	-
	Rural informal (tribal areas)	0.60	0.31	1.15	0.12
	Rural formal (farms)	1.25	0.66	2.37	0.49
Race	Black African	Ref	-	-	-
	White	0.25	0.05	1.14	0.07
	Coloured	0.23	0.08	0.65	0.01
	Indian/Asian	0.01	0.00	0.07	<0.001
Province	Western Cape	Ref	-	-	-
	Eastern Cape	1.08	0.24	4.87	0.92
	Northen Cape	1.15	0.27	4.96	0.85
	Free State	0.65	0.15	2.74	0.56
	KwaZulu-Natal	2.99	0.78	11.46	0.11
	North West	1.16	0.28	4.91	0.84
	Gauteng	1.90	0.45	8.06	0.39
	Mpumalanga	2.92	0.68	12.55	0.15
	Limpopo	0.75	0.16	3.46	0.71
Age		1.09	0.96	1.24	0.18

Table 24.2: Multivariate logistic regression for HIV status in both 2008 and 2017, considering composite risk score and adjusting for socio-demographic variables for females

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk score		1.33	1.18	1.50	<0.001
Survey year	2008	Ref	-	-	-
	2017	0.70	0.53	0.92	0.01
Geographical type	Urban	Ref	-	-	-
	Rural informal (tribal areas)	1.13	0.81	1.57	0.47
	Rural formal (farms)	1.34	0.89	2.02	0.16
Race	Black African	Ref	-	-	-
	White	0.09	0.02	0.42	<0.001
	Coloured	0.30	0.17	0.50	<0.001
	Indian/Asian	0.02	0.01	0.09	<0.001
Province	Western Cape	Ref	-	-	-
	Eastern Cape	1.99	1.03	3.81	0.04
	Northen Cape	1.05	0.46	2.39	0.91
	Free State	0.97	0.48	1.98	0.94
	KwaZulu-Natal	2.15	1.18	3.91	0.01
	North West	0.95	0.48	1.88	0.88
	Gauteng	1.46	0.80	2.67	0.22
	Mpumalanga	2.73	1.50	4.97	<0.001
	Limpopo	0.76	0.37	1.56	0.45
Age		1.26	0.53	0.92	0.01

3.6 Comparison of Results With and Without Multiple Imputation

Multiple imputation is a statistical technique used to handle missing data by creating multiple complete datasets, analysing them separately and then pooling the results to obtain more robust estimates. With multiple imputation, the results for 2008 and 2017 among males and females differed in how missing data influenced estimates (See Appendix for models with missing data).

When stratified by sex in 2008 and 2017, the effect of multiple imputation differed between males and females. Among males, multiple imputation resulted in higher odds ratios with wider confidence intervals for sexual debut and intergenerational sex, and the risk score index,

suggesting greater uncertainty in the imputed estimates. In contrast, among females who had larger proportion of missing data, the multiple imputation led to lower odds ratios for sexual debut and intergenerational sex in the multiple logistic regression model, as well for the risk score index. Additionally, multiple imputation resulted in narrower confidence intervals, indicating increased precision.

This suggests that complete case analysis in females likely overestimated associations due to selection bias, while multiple imputation corrected for this bias and improved precision by re-incorporating excluded cases. MI provided more reliable effect estimates in females, whereas in males the added variability outweighed the benefits. These findings highlight how the impact of MI depends on the extent of missingness, with greater benefits observed in subgroups with higher levels of missing data.

Considering the decomposition analysis for the pooled data, the models generally indicate notable differences in results between data with and without imputation. In the risk predictor model accounting for sexual debut and intergenerational sex among males, the imputed dataset shows narrower confidence intervals, particularly for the overall difference, explained and unexplained components. For the risk score index, the imputed dataset model for females showed a statistically non-significant explained component (Coefficient: -0.001, P = 0.80) while the non-imputed dataset model indicated a statistically significant component (Coefficient: 0.0057, P = 0.01). This suggests that missing data could lead to an overestimation of the contribution of the risk score to the explained difference in HIV prevalence. Once the missing information was accounted for through imputation, the effect was reduced and no longer statistically significant, highlighting the importance of handling missingness appropriately to avoid false associations.

4. DISCUSSION AND CONCLUSION

This study aims to investigate the change in HIV prevalence in young adults aged 15-24 in South Africa between the survey of 2008 and 2017, and to investigate risk factors associated with HIV positivity in both survey years using sex stratified analysis. The study also aims to quantify the separate contributions to the change in HIV prevalence between these surveys, into the portion attributable to change in risk factors and the portion attributable to change in effects of the risk factors, with separate analyses conducted for males and females.

Analysis of young adults aged 15-24 between 2008 and 2017 revealed sex-specific shifts in demographic composition. Among males, the proportion of younger participants (15-19 years) declined, while those aged 20-24 increased, with slight changes in racial composition and

growth in urban and rural formal areas. Provincially, male representation increased in Northern Cape, KwaZulu-Natal, Gauteng, and Mpumalanga, with declines in Eastern Cape, Free State, and Limpopo. Among females, similar age shifts were observed, with increases in Coloured and Indian/Asian representation and decreases in White respondents. Provincially, female representation increased in Western Cape, Northern Cape, Gauteng, and Mpumalanga and decreased in several other provinces. Urban residency grew, while rural areas declined. Consistent urbanization trends occurred across both sexes, likely reflecting migration for education and employment opportunities.

Behavioural risk factors among young adults showed notable sex-specific changes between 2008 and 2017. Among males, early sexual debut, intergenerational sex, transactional sex, and substance use including recreational drugs and excessive alcohol use, became more common. This indicates an increasing engagement in behaviours associated with higher HIV risk. Females also showed rises in intergenerational and transactional sex, as well as substance use, although the increases were generally more modest. Encouragingly, inconsistent condom use declined, particularly among females, suggesting that some HIV prevention efforts, such as condom promotion, may be yielding positive effects. These differences highlight diverging risk profiles by sex. Males appear to be experiencing greater increases in high-risk behaviours, which may contribute to the observed stability or slight rise in HIV prevalence. In contrast, reductions in female HIV prevalence may reflect the combined impact of targeted interventions, improved condom use, and changing behavioural patterns. Overall, these findings underscore the importance of tailored, sex-specific prevention strategies that address both behavioural risks and broader social determinants influencing HIV exposure among young adults.

Contrary to expectations, many high-risk behaviours increased from 2008 to 2017, yet HIV prevalence trends differed by sex. Among males, prevalence showed a slight, non-significant increase, driven in part by higher rates among younger males (15-19 years) and in specific geographic and racial subgroups. In contrast, females experienced a significant decline in HIV prevalence, particularly among older young adults (20-24 years) and Black African females, suggesting that targeted interventions and improvements in prevention practices, such as condom use, may have had a greater impact for females. Age, racial, and geographic disparities were evident for both sexes. While some provinces, such as Gauteng and Mpumalanga, saw substantial declines in female prevalence, others including Eastern Cape and Free State, experienced increases, highlighting regional differences in healthcare access, social determinants, and intervention coverage. Among males, rural formal areas consistently showed the highest prevalence, with urban and rural informal areas experiencing smaller increases, reflecting persistent geographic vulnerabilities. These sex-specific trends

underscore diverging HIV dynamics: males are experiencing stable or slightly increasing prevalence despite rising behavioural risks, whereas females show notable reductions, likely reflecting the effect of targeted interventions. The findings highlight the importance of sex- and region-specific HIV prevention strategies that address behavioural, social, and structural factors influencing HIV risk among young adults.

The sex-specific logistic regression analyses provide important insights into the behavioural risk factors associated with HIV positivity among young adults. Among males, neither early sexual debut nor intergenerational sex was significantly associated with HIV positivity in either survey year, suggesting that these behaviours alone may not be the primary drivers of HIV risk in this group. In contrast, among females, early sexual debut consistently emerged as a strong and significant predictor of HIV positivity. In 2008, females with early sexual debut were 87% more likely to be HIV positive than those with later sexual debut, and in 2017, early sexual debut doubled the odds of HIV positivity compared to later initiation. Intergenerational sex was also associated with higher HIV risk among females, although the association was statistically significant only in 2008. The findings in females align with existing literature indicating the vulnerability of young adults who engage in early sexual activity and intergenerational sex (22,27). These findings indicate clear sex differences in the behavioural determinants of HIV. While early sexual activity and intergenerational relationships remain critical points of intervention for young women, the absence of significant associations among males suggests that other factors may play a more substantial role in determining HIV risk for young men(51,52). Overall, the results underscore the importance of sex-specific HIV prevention strategies that prioritize comprehensive sexual education, delay of sexual debut, and interventions addressing power imbalances in sexual relationships, particularly targeting young women.

Behavioural risk factors associated with HIV positivity were used to derive a risk score index for both 2008 and 2017, with separate analyses for males and females. The score incorporated sexual debut, intergenerational sex, inconsistent condom use, and number of sexual partners, capturing the compounded effect of multiple high-risk behaviours on HIV positivity. In both survey years, the logistic regression models adjusted for socio-demographic variables indicate the risk score was strongly associated with HIV positivity among females. In 2008, each unit increase in the risk score corresponded to a 34% higher likelihood of being HIV positive, and a similar effect was observed in 2017, reinforcing the persistent vulnerability of young women with multiple risk behaviours. These findings imply that the presence of multiple risk behaviours increases vulnerability to HIV infection. Therefore, interventions addressing interlinked behaviours, such as inconsistent condom use and multiple sexual partners are essential (53). Education programs should emphasize the risk associated with

multiple behavioural factors. Among males, however, the risk score was not significantly associated with HIV positivity in either year, suggesting that the cumulative behavioural factors captured by the score may be less predictive of HIV risk in young men. This could indicate that other unmeasured factors such as biological differences, or structural and social determinants (52), may play a more prominent role.

The decomposition analysis by sex is applied to provide insights into the differences in HIV prevalence between the 2008 and 2017 surveys considering sexual debut and intergenerational sex as risk factors. Among males, the explained component showed a non-significant positive contribution to the change in HIV prevalence, with a 0.05% decrease attributable to differences in risk factors over time. Of this, 0.06% was attributed to change in the proportion of individuals participating in early sexual activity, while 0.01% was negatively attributed to the differences in the proportions of individuals participating in intergenerational sex.

The unexplained component among males indicates a non-significant negative contribution to the change of HIV prevalence, with 1.2% of the change attributable to the differences in impact of early sexual debut and intergenerational sex. Of this, 0.33% was attributed to the changes in the impact of sexual debut, while 0.1% was linked to the changes in the impact of intergenerational sex. These findings suggest that the positive explained component, combined with a negative unexplained component, could indicate that less people engaged in these risky behaviours in 2017 as compared to 2008, but the association between these risk factors and HIV positivity has strengthened over time. Therefore, the impact of these risky behaviours over time counteracts the favourable changes in the distribution.

Among females, the explained component indicates a non-significant negative contribution to the change in HIV prevalence, with a reduction of 0.46% attributable to differences in risk factors across the two surveys. Within this component, early sexual debut contributed 0.04%, while intergenerational sex accounted for 0.41%. These results suggest that changes in the distribution of these risk behaviours among females did not substantially drive the observed differences in HIV prevalence over time.

The unexplained component among females, however, indicates a significant positive contribution of 3.42% to the reduction in HIV prevalence between 2008 and 2017. Of this contribution, 0.52% was related to the changed in the impact of early sexual debut, while 0.50% was linked to the changes in the impact of intergenerational sex. These findings suggest that although more females engaged in these risky behaviours over time, the association between these risk factors and HIV positivity weakened over time, leading to favourable changes in HIV prevalence.

Behavioural risk factors that demonstrated a directly proportional association with HIV status were used to derive a risk score index, which was then applied in the decomposition analysis..
Among males, HIV prevalence increased by 1.15% between 2008 and 2017 surveys. Of this, only 0.01% was explained by the differences in the distribution of observed risky behaviours. The explained component revealed that differences among males in observed risky behaviours negatively contributed to a decline in HIV prevalence over time. This indicates that respondents in the 2017 survey had a higher risk score, and therefore, would be expected to have a higher HIV prevalence than respondents in 2008. This highlights the continued need for interventions targeting interlinked risk behaviours such as reducing intergenerational sex, promoting consistent condom use, and minimizing multiple sexual partnerships. The unexplained component in males indicated a non-significant negative contribution in the reduction of HIV prevalence, with an increase of 1.15% attributable to the differences in impact of the risk scores. These findings are consistent with earlier results suggesting that while the prevalence of risky behaviours did not shift substantially, their effect on HIV status was greater in 2017. Overall, this points to a stronger association between risk behaviours and HIV positivity over time, underscoring the importance of addressing both behavioural prevalence and their underlying drivers.
Among females, HIV prevalence decreased by 2.96% between the 2008 and 2017 surveys. Of this change, only 0.1% was explained by differences in the distribution of observed risky behaviours. The explained component showed a negative contribution, suggesting that the prevalence of risky behaviours among females in 2017 would have been expected to increase HIV prevalence compared to 2008. This finding indicates that shifts in behavioural distribution alone did not account for the reduction in HIV prevalence, reinforcing the importance of addressing persistent high-risk patterns such as early sexual debut, inconsistent condom use, and intergenerational sex.
The unexplained component accounted for the majority of the change, with a 3.06% contribution to the reduction in HIV prevalence. This suggests that while the distribution of risk behaviours did not substantially improve, their impact on HIV status weakened over time. Taken together, these findings highlight that reductions in HIV prevalence among females were driven more by changes in the impact of risk factors, rather than by major shifts in behavioural patterns themselves.
A multiple logistic regression analysis was conducted using the pooled survey data from 2008 and 2017, incorporating a survey year variable. Separate models were developed for males and females to provide further insight into the impact of sexual risk factors and later composite risk factors on overall change in HIV prevalence among young adults aged 15-24 years. When accounting for sexual risk factors, the effect of survey year differed by sex. Among males, HIV prevalence was slightly higher in 2017 compared to 2008, though not statistically significant

(aOR 1.12; 95% CI: 0.66–2.21). In contrast, females in 2017 had significantly lower odds of HIV positivity compared to 2008 (aOR 0.68; 95% CI: 0.51–0.90). These patterns are consistent with the decomposition analysis, which showed a non-significant overall increase in males and a significant reduction in HIV prevalence among females between the two survey years.

The results considering the composite risk factors indicate that the effect of survey year differed by sex. Among males, a higher risk score was associated with a lower likelihood of HIV positivity, though this was not statistically significant (aOR 0.90; 95% CI: 0.73–1.12). Males in 2017 were slightly more likely to be HIV-positive compared to 2008, but this increase was not significant (aOR 1.18; 95% CI: 0.64–2.17), suggesting a modest rise in prevalence over time. In contrast, females with higher risk scores were significantly more likely to be HIV-positive (aOR 1.33; 95% CI: 1.18–1.50), and those in 2017 had significantly lower odds of HIV infection compared to 2008 (aOR 0.70; 95% CI: 0.53–0.92), indicating a potential decline in prevalence over time.

In females, the results indicate that the composite risk factors i.e., early sexual debut, intergenerational sex, inconsistent condom use and the number of sexual partners are associated with HIV positivity. These findings also align with the decomposition results, indicating that these combined sexual behaviours are critical factors of HIV prevalence. The significant contribution of the impact in these risk factors to the reduction in HIV prevalence highlight the important associations between risky behaviours and HIV prevalence therefore, emphasizing a need to target these behaviours in prevention strategies.

REFERENCES

1. Fauci AS. Acquired Immunodeficiency Syndrome: Epidemiologic, Clinical, Immunologic, and Therapeutic Considerations. Ann Intern Med. 1984 Jan 1;100(1):92.
2. UNAIDS. Fact Sheet 2024: Global HIV statistics [Internet]. Joint United Nations Programme on HIV/AIDS. 2024 [cited 2024 Aug 8]. Available from: https://www.unaids.org/sites/default/files/media_asset/UNAIDS_FactSheet_en.pdf
3. UNAIDS. UNAIDS DATA 2023 [Internet]. Joint United Nations Programme on HIV/AIDS. 2023 [cited 2024 Aug 12]. Available from: https://www.unaids.org/sites/default/files/media_asset/data-book-2023_en.pdf
4. Joshi K, Lessler J, Olawore O, Loevinsohn G, Bushey S, Tobian AAR, et al. Declining HIV incidence in sub-Saharan Africa: a systematic review and meta-analysis of empiric data. J Int AIDS Soc. 2021;24(10):e25818.
5. UNAIDS. Global Aids update 2018: Miles to Go - Closing Gaps, Breaking Barriers, Righting Injustices. [Internet]. Joint United Nations Programme on HIV/AIDS. 2018

[cited 2024 Aug 7]. Available from: https://www.unaids.org/en/resources/documents/2018/global-aids-update

6. Duby Z, Jonas K, McClinton Appollis T, Maruping K, Dietrich J, Mathews C. "Condoms Are Boring": Navigating Relationship Dynamics, Gendered Power, and Motivations for Condomless Sex Amongst Adolescents and Young People in South Africa. Int J Sex Heal [Internet]. 2021;33(1):40–57. Available from: https://doi.org/10.1080/19317611.2020.1851334

7. Kilburn K, Ranganathan M, Stoner MCD, Hughes JP, MacPhail C, Agyei Y, et al. Transactional sex and incident HIV infection in a cohort of young women from rural South Africa. Aids. 2018;32(12):1669–77.

8. Simelela N, Venter WDF, Pillay Y, Barron P. A Political and Social History of HIV in South Africa. Curr HIV/AIDS Rep. 2015;12:256–61.

9. Chigwedere P, Seage GR, Gruskin S, Lee TH, Essex M. Estimating the Lost Benefits of Antiretroviral Drug Use in South Africa. JAIDS J Acquir Immune Defic Syndr. 2008 Dec 1;49(4):410–5.

10. Rehle TM, Hallett TB, Shisana O, Pillay-van Wyk V, Zuma K, Carrara H, et al. A Decline in New HIV Infections in South Africa: Estimating HIV Incidence from Three National HIV Surveys in 2002, 2005 and 2008. PLoS One. 2010;5(6):e11094.

11. UNAIDS. 90-90-90: An ambitious treatment target to help end the AIDS epidemic. Joint United Nations Programme on HIV/AIDS. 2014.

12. Barron P, Pillay Y, Doherty T, Sherman G, Jackson D, Bhardwaj S, et al. Eliminating mother-to-child HIV transmission in South Africa. Bull World Health Organ. 2013 Jan 1;91(1):70–4.

13. Simelela NP, Venter WDF. A brief history of South Africa's response to AIDS. South African Med J. 2014 Jan 20;104(3):249.

14. Shisana O, Rehle T, Simbayi L, Zuma K, Jooste S, Pillay-Van-Wyk V, et al. South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2008: A turning tide among teenagers? Cape Town: HSRC Press; 2009.

15. Simbayi L, Zuma K, Zungu N, Moyo S, Marinda E, Jooste S, et al. South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2017. Cape Town: HSRC Press; 2019.

16. Peltzer K, Parker W, Mabaso M, Makonko E, Zuma K, Ramlagan S. Impact of national HIV and AIDS communication campaigns in south africa to reduce HIV risk behaviour. Sci World J. 2012;2012.

17. Fapohunda A, Fakolade A, Omiye J, Afolaranmi O, Arowojolu O, Oyebamiji T, et al. Cancer presentation patterns in Lagos, Nigeria: Experience from a private cancer center. J Public Health Africa. 2020 Dec 31;11(2):7.

18. Haffejee F, Maksudi K. Understanding the risk factors for HIV acquisition among refugee women in South Africa. AIDS Care. 2020 Jan 2;32(1):37–42.

19. Stoner MCD, Pettifor A, Edwards JK, Aiello AE, Halpern CT, Julien A, et al. The effect of school attendance and school dropout on incident HIV and HSV-2 among young women in rural South Africa enrolled in HPTN 068. Aids. 2017;31(15):2127–34.

20. HSRC. New HIV Survey Highlights Progress and Ongoing Disparities in South Africa's HIV Epidemic [Internet]. 2023 [cited 2024 Oct 31]. Available from: https://hsrc.ac.za/press-releases/phsb/new-hiv-survey-highlights-progress-and-

ongoing-disparities-in-south-africas-hiv-epidemic/

21. Wabiri N, Taffa N. Socio-economic inequality and HIV in South Africa. BMC Public Health. 2013 Dec 4;13(1):1037.

22. Bajunirwe F, Semakula D, Izudi J. Risk of HIV infection among adolescent girls and young women in age-disparate relationships in sub-Saharan Africa. AIDS. 2020 Aug 1;34(10):1539–48.

23. Lewis L, Kharsany ABM, Humphries H, Maughan-Brown B, Beckett S, Govender K, et al. HIV incidence and associated risk factors in adolescent girls and young women in South Africa: A population-based cohort study. PLoS One. 2022;17(12).

24. Weller SC, Davis-Beaty K. Condom effectiveness in reducing heterosexual HIV transmission. Cochrane Database Syst Rev. 2002 Jan 21;2012(1).

25. Morris M, Kretzschmar M. Concurrent partnerships and the spread of HIV. AIDS. 1997;11(5):641–8.

26. Heeren AG, Mandeya A, Jemmott JB, Chiruka RT, Marange SC, Batidzirai JM, et al. Multiple partners and condom use among students at a South African University. J Evid Based Soc Work. 2014;11(5):437–44.

27. Shisana O, Rehle T, Simbayi L, Zuma K, Jooste S, Zungu N, et al. South African National HIV Prevalence, Incidence & Behaviour Survey, 2012. Cape Town: HSRC Press; 2014.

28. Nigatu AM, Birhanu AY, Endehabtu BF, Zimmerman L, Olson H, Tsui A, et al. Early Coital Debut and Associated HIV Risk Factors Among Young Women and Men in South Africa. Int Perspect Sex Reprod Health. 2009;35(2):82–90.

29. Qian H zhu, Schumacher JE, Chen HT, Ruan Y hua. Injection drug use and HIV/AIDS in China: Review of current situation, prevention and policy implications. Harm Reduct J. 2006;3(1).

30. Abbey A. Alcohol's role in sexual violence perpetration: Theoretical explanations, existing evidence and future directions. Drug Alcohol Rev. 2011;30(5):481–9.

31. Jahagirdar D, Walters M, Vongpradith A, Dai X, Novotney A, Kyu HH, et al. Incidence of HIV in Sub-Saharan Africa, 2000–2015: The Interplay Between Social Determinants and Behavioral Risk Factors. AIDS Behav. 2021 Nov 5;25(S2):145–54.

32. Chun HM, Carpenter RJ, Macalino GE, Crum-Cianflone NF. The Role of Sexually Transmitted Infections in HIV-1 Progression: A Comprehensive Review of the Literature. J Sex Transm Dis. 2013;2013.

33. Kalichman SC, Pellowski J, Turner C. Prevalence of sexually transmitted co-infections in people living with HIV/AIDS: systematic review with implications for using HIV treatments for prevention. Sex Transm Infect. 2011 Apr 1;87:183–90.

34. N WC, A S. Associated Risk Factors of STIs and Multiple Sexual Relationships among Youths in Malawi. Dalal K, editor. PLoS One. 2015 Aug 6;10(8):e0134286.

35. ScienceDaily. How STDs Increase the Risk of Becoming Infected with HIV. J Clin Invest [Internet]. 2008 [cited 2024 Nov 2]; Available from: www.sciencedaily.com/releases/2008/09/080905215948.htm

36. Lavoie MCC, Blanco N, Lascko T, Baumhart C, Verbeek J, Stoebenau K, et al. The effects of the Determined, Resilient, Empowered, AIDS-free, Mentored, and Safe (DREAMS) program on HIV incidence and other health-related outcomes among adolescent girls and young women: a systematic review and meta-analysis.

eClinicalMedicine [Internet]. 2025;82:103176. Available from: https://doi.org/10.1016/j.eclinm.2025.103176

37. Govender K, Beckett S, Reddy T, Cowden RG, Cawood C, Khanyile D, et al. Association of HIV Intervention Uptake With HIV Prevalence in Adolescent Girls and Young Women in South Africa. JAMA Netw Open. 2022;5(4).

38. UNAIDS. Undetectable = Untransmittable: Public Health and HIV Viral Load Suppression [Internet]. UNAIDS Explainer. 2024. Available from: https://www.unaids.org/en/resources/documents/2018/undetectable-untransmittable

39. Lilian RR, Rees K, McIntyre JA, Struthers HE, Peters RPH. Same-day antiretroviral therapy initiation for HIV-infected adults in South Africa: Analysis of routine data. PLoS One [Internet]. 2020;15(1):1–13. Available from: http://dx.doi.org/10.1371/journal.pone.0227572

40. Blinder AS. Wage Discrimination: Reduced Form and Structural Estimates. J Hum Resour. 1973;8(4):436–55.

41. Fairlie RW. An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. J Econ Soc Meas. 2005 Dec 29;30(4):305–16.

42. Jann B. The Blinder–Oaxaca Decomposition for Linear Regression Models. Stata J Promot Commun Stat Stata. 2008 Dec 1;8(4):453–79.

43. Zaba B, Boerma T, White R. Monitoring the AIDS epidemic using HIV prevalence data among young women attending antenatal clinics: prospects and problems. AIDS. 2000 Jul;14(11):1633–45.

44. The International Group on Analysis of Trends in HIV Prevalence and Behaviours in Young People in Countries most Affected by HIV. Trends in HIV prevalence and sexual behaviour among young people aged 15–24 years in countries most affected by HIV. Sex Transm Infect. 2010;86(Suppl 2):ii72–83.

45. Donovan B, Ross MW. Preventing HIV: determinants of sexual behaviour. Lancet. 2000;355(9218):1897–901.

46. Blinder AS. Wage Discrimination: Reduced Form and Structural Estimates. J Hum Resour. 1973;8(4):436–55.

47. Oaxaca R. Male-Female Wage Differentials in Urban Labor Markets Author. Int Econ Rev (Philadelphia). 1973;14(3):693–709.

48. Fairlie, Robert. W (University of Carlifornia SC. The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment. J Labor Econ. 1999;17(1):80–108.

49. Yun MS. Decomposing differences in the first moment. Econ Lett. 2004;82(2):275–80.

50. Oaxaca RL, Ransom MR. Identification in Detailed Wage Decompositions. Rev Econ Stat. 1999 Feb 1;81(1):154–7.

51. Baral S, Logie C, Grosso A, Wirtz AL, Beyrer C. Modified Social Ecological Model: A tool to visually represent the risks and risk contexts of HIV epidemics. BMC Public Health [Internet]. 2013;13:482. Available from: http://www.biomedcentral.com/1471-2458/13/482

52. Stannah J, Flores Anato JL, Pickles M, Larmarange J, Mitchell KM, Artenie A, et al. From conceptualising to modelling structural determinants and interventions in HIV transmission dynamics models: a scoping review and methodological framework for evidence-based analyses. BMC Med [Internet]. 2024;22(1). Available from:

https://doi.org/10.1186/s12916-024-03580-z

53. Shahmanesh M, Patel V, Mabey D, Cowan F. Effectiveness of interventions for the prevention of HIV and other sexually transmitted infections in female sex workers in resource poor setting: A systematic review. Trop Med Int Heal. 2008;13(5):659–79.

APPENDIX

Appendix A: Risk Factors Associated with HIV Positivity in 2008 Adjusting for Socio-Demographic Variables (Data with Missing Values, No Imputation)

Logistic regression models for HIV status in 2008 adjusting for socio-demographic variables in males

Variable		**Odds ratio**	**95% Conf. Interval**		**P - value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<= 16yrs)	0.85	0.34	2.14	0.73
Intergenerational sex	No	**Ref**	-	-	-
	Yes	1.19	0.27	5.21	0.81
Consistent condom use	Not at risk	**Ref**	-	-	-
	At risk	0.90	0.26	3.17	0.87
	No	**Ref**	-	-	-

Recreational drug use	Yes	1.57	0.31	8.09	0.59
Excessive alcohol	No	**Ref**	-	-	-
	Yes	1.18	0.38	3.67	0.77
Number of sexual partners		1.00	0.80	1.25	1.00

Logistic regression models for HIV status in 2008 adjusting for socio-demographic variables in females

Variable		**Odds ratio**	**95% Conf. Interval**		**P - value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<= 16yrs)	2.32	1.52	3.53	0.00
Intergenerational sex	No	**Ref**	-	-	-
	Yes	2.23	1.37	3.64	0.00
Consistent condom use	Not at risk	**Ref**	-	-	-
	At risk	1.10	0.66	1.84	0.72
Transactional sex	No	**Ref**	-	-	-
	Yes	0.76	0.33	17.45	0.87
Recreational drug use	No	**Ref**	-	-	-
	Yes	1.27	0.22	7.46	0.79
Excessive alcohol	No	**Ref**	-	-	-
	Yes	1.47	0.69	3.12	0.31
Number of sexual partners		1.04	0.90	1.20	0.61

Multivariate logistic regression for HIV status in 2008 considering predictors and adjusting for socio-demographic variables in males

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<=16yrs)	0.77	0.28	2.16	0.62
Intergenerational sex	No	**Ref**	-	-	-
	Yes	1.21	0.25	5.91	0.82
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.27	0.07	1.13	0.07
	Rural formal (farms)	0.26	0.03	2.35	0.23
Race	Black African	**Ref**	-	-	-
	White	1			
	Coloured	0.03	0.00	0.28	<0.001

	Indian/Asian	1			
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.50	0.12	18.40	0.75
	Northen Cape	4.53	0.48	42.82	0.19
	Free State	1.44	0.21	9.95	0.71
	KwaZulu-Natal	11.32	2.10	61.10	0.01
	North West	4.46	0.68	29.04	0.12
	Gauteng	7.01	1.23	45.79	0.04
	Mpumalanga	11.40	1.23	105.75	0.03
	Limpopo	0.66	0.06	7.22	0.73
Age		1.05	0.80	1.38	0.71

Multivariate logistic regression for HIV status in 2008 considering predictors and adjusting for socio-demographic variables in females

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<=16yrs)	2.14	1.37	3.36	<0.001
Intergenerational sex	No	**Ref**	-	-	-
	Yes	2.02	1.22	3.35	0.01
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	1.02	0.60	1.74	0.95
	Rural formal (farms)	1.51	0.77	2.96	0.23
Race	Black African	**Ref**	-	-	-
	White	0.01	0.00	0.06	<0.001
	Coloured	0.19	0.08	0.41	<0.001
	Indian/Asian	0.00	0.00	0.01	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.59	0.55	4.60	0.40
	Northen Cape	1.07	0.31	3.72	0.91
	Free State	0.60	0.19	1.87	0.37
	KwaZulu-Natal	2.86	1.09	7.54	0.03
	North West	0.90	0.29	2.78	0.86
	Gauteng	1.93	0.73	5.11	0.19
	Mpumalanga	2.32	0.82	6.56	0.11

	Limpopo	0.75	0.22	2.61	0.65
Age		1.31	1.20	1.44	<0.001

Logistic regression for HIV status in 2008 considering risk score index and adjusting for socio-demographic variables in males

Variables		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk Score index		1.04	0.76	1.42	0.80
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.32	0.08	1.29	0.11
	Rural formal (farms)	0.28	1.29	2.58	0.26
Race	Black African	**Ref**	-	-	-
	White	1			
	Coloured	0.03	0.00	0.29	<0.001
	Indian/Asian	1			
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.61	0.13	19.76	0.71
	Northen Cape	5.24	0.55	50.04	0.15
	Free State	1.58	0.23	10.95	0.64
	KwaZulu-Natal	11.40	2.09	62.13	0.01
	North West	5.16	0.82	32.61	0.08
	Gauteng	7.65	1.18	49.62	0.03
	Mpumalanga	11.86	1.47	95.90	0.02
	Limpopo	0.68	0.06	7.25	0.74
Age		1.02	0.79	1.31	0.87

Logistic regression for HIV status in 2008 considering risk score index and adjusting for socio-demographic variables in females

Variables		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk Score index		1.45	1.20	1.75	<0.001
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.98	0.59	1.63	0.93
	Rural formal (farms)	1.48	0.79	2.80	0.22
Race	Black African	**Ref**	-	-	-
	White	0.01	0.00	0.06	<0.001
	Coloured	0.18	0.08	0.40	<0.001
	Indian/Asian	0.00	0.00	0.01	<0.001
Province	Western Cape	**Ref**	-	-	-

	Eastern Cape	1.63	0.57	4.67	0.36
	Northen Cape	0.98	0.30	3.21	0.97
	Free State	0.64	0.20	2.00	0.44
	KwaZulu-Natal	2.93	1.14	7.53	0.03
	North West	0.86	0.28	2.60	0.78
	Gauteng	1.81	0.69	4.75	0.23
	Mpumalanga	2.23	0.80	6.19	0.12
	Limpopo	0.68	0.20	2.33	0.54
Age		1.29	1.18	1.41	<0.001

Appendix B: Risk Factors Associated with HIV Positivity in 2017 Adjusting for Socio-Demographic Variables (Data with Missing Values, No Imputation)

Logistic regression for HIV status in 2017 adjusting for socio-demographic variables in males

Variable		**Odds ratio**	**95% Conf. Interval**		**P - value**
Sexual Debut	Not at Risk	**Ref**	-	-	-
	Risk (<= 16yrs)	0.62	0.33	1.15	0.13
Intergenerational sex	No	**Ref**	-	-	-
	Yes	0.90	0.34	2.39	0.83
Consistent condom use	Not at risk	**Ref**	-	-	-
	At risk	1.82	0.68	4.83	0.23
Transactional sex	No	**Ref**	-	-	-
	Yes	0.55	0.08	3.57	0.53
Recreational drug use	No	**Ref**	-	-	-
	Yes	0.49	0.19	1.27	0.14
Excessive alcohol	No	**Ref**	-	-	-
	Yes	1.24	0.54	2.88	0.61
Number of sexual partners		0.71	0.45	1.10	0.13

Logistic regression for HIV status in 2017 adjusting for socio-demographic variables in females

Variable		**Odds ratio**	**95% Conf. Interval**		**P - value**
Sexual Debut	Not at Risk	**Ref**	-	-	-
	Risk (<= 16yrs)	2.41	1.72	3.37	<0.001
Intergenerational sex	No	**Ref**	-	-	-
	Yes	1.62	1.03	2.53	0.04
Consistent condom use	Not at risk	**Ref**	-	-	-
	At risk	1.18	0.69	2.03	0.55
Transactional sex	No	**Ref**	-	-	-
	Yes	1.96	0.56	6.90	0.29
Recreational drug use	No	**Ref**	-	-	-
	Yes	1.41	0.64	3.10	0.40
Excessive alcohol	No	**Ref**	-	-	-
	Yes	0.76	0.47	1.24	0.27
Number of sexual partners		1.42	1.17	1.80	<0.001

Logistic regression for HIV status in 2017 considering predictors and adjusting for socio-demographic variables in males

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<=16yrs)	0.58	0.30	1.12	0.10
Intergenerational sex	No	Ref	-	-	-
	Yes	1.02	0.38	2.76	0.97
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.64	0.33	1.22	0.17
	Rural formal (farms)	2.47	1.19	5.10	0.01
Race	Black African	**Ref**	-	-	-
	White	0.19	0.03	1.06	0.06
	Coloured	0.17	0.04	0.79	0.02
	Indian/Asian	0.03	0.00	0.24	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.17	0.20	6.83	0.86
	Northen Cape	0.41	0.07	2.51	0.34
	Free State	0.44	0.08	2.55	0.36
	KwaZulu-Natal	1.12	0.23	5.53	0.89
	North West	0.72	0.13	4.13	0.71
	Gauteng	0.87	0.16	4.72	0.87
	Mpumalanga	1.07	0.20	5.73	0.94
	Limpopo	0.70	0.12	4.08	0.70
Age		1.06	0.92	1.21	0.42

Logistic regression for HIV status in 2017 considering predictors and adjusting for socio-demographic variables in females

Variable		Odds ratio	95% Conf. Interval		P-value
Sexual Debut	Not at risk	**Ref**	-	-	-
	Risk (<=16yrs)	2.25	1.58	3.20	<0.001
Intergenerational sex		**Ref**	-	-	-
		1.36	0.85	2.16	0.20
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	1.16	0.77	1.74	0.48
	Rural formal (farms)	1.34	0.76	2.38	0.31
Race	Black African	**Ref**	-	-	-
	White	0.05	0.02	0.18	<0.001
	Coloured	0.38	0.18	0.77	<0.001
	Indian/Asian	0.10	0.02	0.45	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.89	0.93	3.85	0.08
	Northen Cape	0.57	0.22	1.50	0.25
	Free State	1.12	0.47	2.65	0.80
	KwaZulu-Natal	1.23	0.60	2.51	0.57
	North West	0.91	0.42	1.97	0.81
	Gauteng	0.99	0.47	2.09	0.98
	Mpumalanga	1.80	0.88	3.69	0.11
	Limpopo	0.75	0.33	1.70	0.49
Age		1.21	1.13	1.30	<0.001

Logistic regression for HIV status in 2017 considering risk score index and adjusting for socio-demographic variables in males

Variables		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk Score		0.80	0.61	1.06	0.12
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	0.68	0.37	1.27	0.23
	Rural formal (farms)	2.34	1.12	4.89	0.02
Race	Black African	**Ref**	-	-	-
	White	0.20	0.04	1.12	0.07
	Coloured	0.16	0.03	0.80	0.03
	Indian/Asian	0.03	0.00	0.25	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.13	0.18	6.98	0.90
	Northen Cape	0.41	0.07	2.52	0.33
	Free State	0.43	0.07	2.66	0.36
	KwaZulu-Natal	1.04	0.20	5.34	0.97
	North West	0.70	0.12	4.24	0.70
	Gauteng	0.83	0.14	4.75	0.83
	Mpumalanga	1.18	0.21	6.64	0.85
	Limpopo	0.77	0.13	4.56	0.78
Age		1.08	0.94	1.22	0.27

Logistic regression for HIV status in 2017 considering risk score index and adjusting for socio-demographic variables in males

Variables	**Odds ratio**	**95% Conf. Interval**	**P-value**

Risk Score		1.37	1.18	1.60	<0.001
Geographical type	Urban	**Ref**	-	-	-
	Rural informal (tribal areas)	1.29	0.88	1.91	0.19
	Rural formal (farms)	1.51	0.85	2.66	0.16
Race	Black African	**Ref**	-	-	-
	White	0.06	0.02	0.20	<0.001
	Coloured	0.36	0.17	0.74	0.01
	Indian/Asian	0.10	0.02	0.44	<0.001
Province	Western Cape	**Ref**	-	-	-
	Eastern Cape	1.91	0.95	3.84	0.07
	Northen Cape	0.60	0.23	1.56	0.29
	Free State	1.15	0.49	2.66	0.75
	KwaZulu-Natal	1.21	0.61	2.41	0.59
	North West	0.82	0.38	1.76	0.61
	Gauteng	0.97	0.47	1.98	0.93
	Mpumalanga	1.71	0.84	3.46	0.14
	Limpopo	0.64	0.29	1.44	0.28
Age		1.18	1.10	1.27	<0.001

Appendix C: Risk factors associated with HIV positivity for pooled survey data. (Data with Missing Values, No Imputation)

Decomposition analysis of HIV status using risk predictors in males

Decomposition analysis		Coefficient	95% Conf. Interval		P - value
Overall	Difference	- 0.0113	- 0.035	0.013	0.36
	Explained	0.0010	- 0.002	0.003	0.44
	Unexplained	- 0.0123	- 0.04	0.011	0.30
Explained	Sexual Debut	0.0010	- 0.001	0.004	0.42
	Intergenerational sex	- 0.0001	- 0.000	0.000	0.72
Unexplained	Sexual Debut	0.0032	- 0.012	0.019	0.69
	Intergenerational sex	0.0005	- 0.003	0.004	0.77

Decomposition analysis of HIV status using risk predictors in females

Decomposition analysis		Coefficient	95% Conf. Interval		P - value
Overall	Difference	0.0298	0.001	0.0590	0.046
	Explained	- 0.0046	- 0.009	0.000	0.060
	Unexplained	0.0344	0.005	0.064	0.022
Explained	Sexual Debut	- 0.0004	- 0.004	0.003	0.82
	Intergenerational sex	- 0.0042	- 0.008	- 0.001	0.02
Unexplained	Sexual Debut	- 0.0047	- 0.019	0.009	0.51
	Intergenerational sex	0.0054	- 0.007	0.018	0.40

Decomposition analysis of HIV status using risk score index in males

Decomposition analysis		Coefficient	95% Conf. Interval		P - Value
Overall	Difference	- 0.0115	- 0.034	0.011	0.31
	Explained	0.0002	- 0.001	0.001	0.66
	Unexplained	- 0.0118	- 0.034	0.010	0.29

Explained	Risk score	0.0002	- 0.001	0.001	0.66
Unexplained	Risk score	0.006	- 0.010	0.022	0.48

Decomposition analysis of HIV status using risk score index in females

Decomposition analysis		**Coefficient**	**95% Conf. Interval**		**P - Value**
Overall	Difference	0.0296	0.001	0.059	0.04
	Explained	0.0057	0.001	0.010	0.01
	Unexplained	0.0239	- 0.004	0.052	0.09
Explained	Risk score	0.0057	0.001	0.010	0.01
Unexplained	Risk score	0.0012	- 0.019	0.021	0.91

Multivariate logistic regression for HIV status in both 2008 and 2017, considering predictors and adjusting for socio-demographic variables in males

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	Ref	-	-	-
	Risk (<=16yrs)	0.61	0.34	1.10	0.10
Intergenerational	No	Ref	-	-	-

sex	Yes	1.15	0.49	2.73	0.75
Survey year	2008	Ref	-	-	-
	2017	1.17	0.60	2.27	0.64
Geographical type	Urban	Ref	-	-	-
	Rural informal (tribal areas)	0.42	0.21	0.82	0.01
	Rural formal (farms)	1.21	0.58	2.50	0.61
Race	Black African	Ref	-	-	-
	White	0.07	0.01	0.38	<0.001
	Coloured	0.12	0.03	0.43	<0.001
	Indian/Asian	0.01	0.00	0.07	<0.001
Province	Western Cape	Ref	-	-	-
	Eastern Cape	1.09	0.23	5.16	0.91
	Northen Cape	0.94	0.19	4.73	0.94
	Free State	0.54	0.12	2.39	0.41
	KwaZulu-Natal	2.46	0.60	10.05	0.21
	North West	1.14	0.26	5.00	0.86
	Gauteng	1.64	0.36	7.45	0.52
	Mpumalanga	2.28	0.47	11.05	0.31
	Limpopo	0.72	0.14	3.58	0.69
Age		1.06	0.92	1.21	0.42

Multivariate logistic regression for HIV status in both 2008 and 2017, considering predictors and adjusting for socio-demographic variables in females

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Sexual Debut	Not at risk	Ref	-	-	-
	Risk (<=16yrs)	2.15	1.61	2.87	0.00
Intergenerational sex	No	Ref	-	-	-
	Yes	1.66	1.18	2.33	0.00
Survey year	2008	Ref	-	-	-
	2017	0.67	0.51	0.88	0.00
	Urban	Ref	-	-	-

Geographical type	Rural informal (tribal areas)	1.04	0.73	1.47	0.84
	Rural formal (farms)	1.38	0.89	2.15	0.15
Race	Black African	Ref	-	-	-
	White	0.02	0.01	0.06	0.00
	Coloured	0.28	0.16	0.49	0.00
	Indian/Asian	0.03	0.01	0.12	0.00
Province	Western Cape	Ref	-	-	-
	Eastern Cape	1.67	0.87	3.23	0.12
	Northen Cape	0.81	0.37	1.79	0.60
	Free State	0.84	0.40	1.73	0.63
	KwaZulu-Natal	2.08	1.13	3.83	0.02
	North West	0.92	0.46	1.86	0.83
	Gauteng	1.44	0.78	2.65	0.25
	Mpumalanga	2.13	1.14	3.99	0.02
	Limpopo	0.78	0.36	1.68	0.53
Age		1.27	1.19	1.34	0.00

Multivariate logistic regression for HIV status in both 2008 and 2017, considering composite risk score and adjusting for socio-demographic variables in males

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk score		0.87	0.70	1.07	0.18
Survey year	2008	Ref	-	-	-
	2017	1.17	0.64	2.16	0.60
Geographical type	Urban	Ref	-	-	-
	Rural informal (tribal areas)	0.59	0.32	1.11	0.10
	Rural formal (farms)	1.22	0.65	2.30	0.53
Race	Black African	Ref	-	-	-
	White	0.24	0.05	1.12	0.07
	Coloured	0.23	0.08	0.64	0.01
	Indian/Asian	0.01	0.00	0.06	0.00

Province	Western Cape	Ref	-	-	-
	Eastern Cape	1.07	0.24	4.85	0.93
	Northen Cape	1.14	0.26	4.91	0.86
	Free State	0.66	0.15	2.77	0.57
	KwaZulu-Natal	2.96	0.77	11.33	0.11
	North West	1.16	0.27	4.93	0.84
	Gauteng	1.90	0.44	8.12	0.39
	Mpumalanga	2.87	0.67	12.31	0.16
	Limpopo	0.75	0.16	3.47	0.72
Age		1.09	0.96	1.23	0.18

Multivariate logistic regression for HIV status in both 2008 and 2017, considering composite risk score and adjusting for socio-demographic variables in females

Variable		**Odds ratio**	**95% Conf. Interval**		**P-value**
Risk score		1.34	1.19	1.50	0.00
Survey year	2008	Ref	-	-	-
	2017	0.73	0.56	0.96	0.03
Geographical type	Urban	Ref	-	-	-
	Rural informal (tribal areas)	1.17	0.84	1.62	0.36
	Rural formal (farms)	1.44	0.97	2.13	0.07
Race	Black African	Ref	-	-	-
	White	0.09	0.02	0.46	0.00
	Coloured	0.29	0.17	0.50	0.00
	Indian/Asian	0.02	0.01	0.09	0.00
Province	Western Cape	Ref	-	-	-
	Eastern Cape	2.05	1.08	3.92	0.03
	Northen Cape	1.13	0.50	2.53	0.77
	Free State	0.94	0.47	1.91	0.87
	KwaZulu-Natal	2.18	1.22	3.92	0.01
	North West	0.92	0.47	1.80	0.80

	Gauteng	1.44	0.80	2.61	0.23
	Mpumalanga	2.77	1.51	5.06	0.00
	Limpopo	0.74	0.36	1.53	0.42
Age		1.27	1.20	1.35	0.00

Plagiarism

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

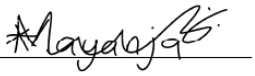
PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I Moses Micah Mayanja (Student number: 2388532) am a student registered for the degree of Master of Science in Feild Biostatistics in the academic year 2022.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature:  Date: 26 - March - 2025

Turnitin cover page

Moses Micah Mayanja - Research Report-1.docx

Signed

Prof Jonathan Levin
Supervisor

ORIGINALITY REPORT

20%	15%	14%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Witwatersrand Student Paper	2%
2	hdl.handle.net Internet Source	1%
3	Ouma, Joseph Douglas Oguti. "Monitoring HIV/AIDS Dynamics: An Application of Real Time Assessment Methods for Estimating Prevalence at District Levels in Uganda", University of the Witwatersrand, Johannesburg (South Africa), 2025 Publication	1%
4	"Encyclopedia of AIDS", Springer Science and Business Media LLC, 2018 Publication	1%
5	www.hsrc.ac.za Internet Source	1%
6	www.saldru.uct.ac.za Internet Source	1%
7	link.springer.com Internet Source	<1%
8	saspublishers.com Internet Source	<1%
9	Submitted to London School of Hygiene and Tropical Medicine Student Paper	<1%

Ethics waiver

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG	HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

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

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
ETHICS WAIVER

27/09/2024

Ref: W-MvN-240927-01

TO WHOM IT MAY CONCERN

Waiver: This certifies that the following research does not require clearance from the Human Research Ethics Committee (Medical)

Investigator: Mr MM Mayanja
Student No. (if appropriate): 2388532
Staff No. (if appropriate):

Supervisor: Professor J Levin

School: Public Health
Department:
Division: Epidemiology and Biostatistics
Faculty of Health Sciences

Project title: *Trends in HIV prevalence and risk factors amongst young adults aged 15-24 years in South Africa, 2008-2017*

Degree: MSc

Reason: Review of information in the public domain
No human participants will be involved in the study

Ms M van Niekerk
Co-Chairperson: Human Research Ethics Committee (Medical)

Research Office Secretariat:
Third Floor, Phillip Tobias Building, corner of St Andrews and York Roads, Parktown, Johannesburg 2193
Postal address: Private Bag 3, Wits 2050
Tel Nos: +27 (0)11 717 1234/1252/2656/2700
Office E-mail: HREC-Medical.ResearchOffice@wits.ac.za
Website:
https://www.wits.ac.za/research/researcher-support/research-ethics/ethics-committees/