# The relationship between individual and neighbourhood socioeconomic factors and HIV prevalence in a national population based survey conducted in Zambia.

**Chola Nakazwe**

**Centre for International Health**

**Faculty of Medicine and Dentistry**

**University of Bergen, Norway**

**2012**

# The relationship between individual and neighbourhood socioeconomic factors and HIV prevalence in a national population based survey conducted in Zambia.

**Chola Nakazwe**

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Philosophy in International Health at the University of Bergen.

Centre for International Health

Faculty of Medicine and Dentistry

University of Bergen, Norway 2012

# ABSTRACT

**Background**: Emerging issues in HIV prevention include the importance of considering underlying social and economic factors at the community and individual level. We examined the associations between individual and neighbourhood socioeconomic position (SEP) on HIV prevalence in young people in a high HIV prevalence country.

**Methods**: The study re-analysed data from the Zambia Demographic and Health Survey, a cross-sectional nationally representative survey conducted in 2007. A two stage cluster stratified systematic sampling procedure was used to select a sample of 14,554, of which 10,444 were successfully interviewed and tested for HIV. Young people (aged 15-24 years) accounted for 40.7% of the sample (n=4,253). Neighbourhood level variables were derived by aggregating the individual level SEP indicators wealth, educational attainment and employment. A composite SEP indicator was constructed by a linear combination of SEP indicators and a neighbourhood level SEP indicator was derived by aggregating the composite SEP indictor. Multi-level mixed effects logistic regression models were used to examine the association between HIV prevalence and different measures of SEP at individual and neighbourhood level.

**Results:** HIV prevalence among young people was 6.5%. The prevalence was higher in urban than rural areas (8.5% compared to 4.7%). A higher proportion of young women (8.5%) were infected with HIV compared to young men (4.3%). In rural areas, young people from high employment neighbourhoods were less likely to be infected with HIV (Age adjusted odds ratio (AOR) 0.34, 95% CI 0.18 – 0.63). However, living in neighbourhoods with middle educational attainment and the wealthiest neighbourhoods increased the risk of HIV infection. Neighbourhoods that scored high on the aggregate composite SEP indicator were also associated with increased odds of infection. At individual level better education and high composite SEP reduced the odds of HIV infection in urban areas, whereas there was no significant association for these two individual variables in rural areas. Employment reduced the odds of being infected in both rural and urban areas ((AOR 0.53, 95% CI 0.34 – 0.81) and (AOR 0.65, 95% CI 0.46 – 0.91), respectively). Controlling for neighbourhood SEP indicators did not affect the significant association of individual SEP indicators in urban areas, while in rural areas the associations were affected.

**Conclusion**: There were marked differentials by residence in the way underlying socioeconomic factors affects HIV transmission both at the individual and community level. Our results suggest that community level factors have a more important influence in rural than urban areas. Preventive strategies targeting community level factors are urgently needed.

# TABLE OF CONTENTS

# LIST OF TABLES AND FIGURES

**Tables**



**Figures**

# Acronyms and Abbreviations

| | | |
|---|---|---|
| AIDS | - | Acquired Immunodeficiency Syndrome |
| ANC | - | Antenatal Care |
| CDC | - | Centre for Disease Control and Prevention |
| CSO | - | Central Statistical Office |
| HIV | - | Human Immunodeficiency Virus |
| MOH | - | Ministry of Health |
| MOE | - | Ministry of Education |
| SEP | - | Socioeconomic Position |
| SES | - | Socioeconomic Status |
| STI | - | Sexually Transmitted Infections |
| TDRC | - | Tropical Disease and Research Centre |
| UNAIDS | - | Joint United Nations Programme on HIV/AIDS |
| VCT | - | Voluntary Counselling and Testing |
| WHO | - | World Health Organization |
| ZDHS | - | Zambia Demographic and Health Survey |

# ACKNOWLEDGEMENTS

I would like to acknowledge the Norwegian government for their financial support through the Quota scholarship scheme for my master's research work and studies at University of Bergen. I would also like to acknowledge the support of the management at Central Statistical Office and University of Zambia (Community Medicine).

My sincere appreciation goes to my Professor Knut Martin Fylkesnes for his dedicated guidance, technical advice and sharing his valuable experience during my studies. My appreciation also goes to my co-supervisor Ingvild Fossgard Sandøy for her guidance, advice and meticulous editing. Both my supervisors made the research learning experience a joyous experience for me.

I extend my heartfelt gratitude to my loving husband and children for their love, support and understanding. I love you so much and you mean the world to me. To my brother and sisters (Webster, Miya, Chits, Yasa and Viv), thank you for your encouragement and love. To Pastor Bruce, Pastor Hannah, Dr. Liya and my Mt. Zion family, I am grateful for your spiritual guidance and prayers. To the Daka's, Aunty Esther, Carol, Getrude and Eliza, thank you taking care of my family and your support. Lister and Lumbwe, I appreciate your love and concern for me. You made my stay in Bergen bearable and fun. I am grateful to Nkombo for sharing his research experience and advice. Thank you to my workmates and friends, Dorothy, Sheila, Josephine, Nchimunya, Margaret, Namakao and Patricia. And my new friends in Bergen, Selia, Linda, Mpundu and Mark, just to mention a few. Lastly my appreciation goes to all my friends and the staff at CIH.

Above all I would like to thank God for His love, grace and strength throughout my studies.

# 1.0. INTRODUCTION

## 1.1. OVERVIEW/HIV SITUATION

The HIV epidemic remains a major concern on the global health agenda. UNAIDS estimates that globally 33.3 million people are infected with Human Immunodeficiency Virus (HIV) and Sub-Saharan African bears a huge burden of the HIV infections. The region accounts for over two thirds (68%) of the global total. The HIV prevalence is estimated to be 5% in Sub-Saharan Africa on average. However, there are considerable regional variations in the epidemics, with Southern Africa the most severely affected. Countries with the highest prevalence in the world are in Southern Africa, with prevalence as high as 25.9% in Swaziland [1].

## 1.2. YOUNG PEOPLE AND HIV

Young people are an important target group in the fight against HIV and AIDS. They are at a stage when sexual and reproductive activity including risky sexual behaviour are likely to begin, therefore averting infections in this age group is critical in the global response to the epidemic [2]. One of the key indicators of the ability of National AIDS control programmes to control the epidemic is HIV prevalence among young people aged 15-24 years [3]. Prevalence estimates restricted to young people are more likely to reflect new infections and are also not likely to be affected by mortality or AIDS treatment. The UNAIDS estimates that 3.8 .million young people in Sub-Saharan Africa are living with HIV, accounting for 79% of the global infections in young people [4]. Current data shows a changing positive pattern in some of the most severely affected countries (Botswana, South Africa, Zambia and Zimbabwe). Statistically significant declines in prevalence were observed among young women and men in national surveys [1]. In Zambia 6.5% of the young people were HIV positive. HIV prevalence is higher for young women (9%) than men (4%). A linear increase of HIV prevalence with age was observed for women, from 4% for those aged 15-17, to 6% for those aged 18-22 years then to 13% for the age 23-24 years. The pattern was not liner for men [5]. In the absence of incidence data, studying HIV epidemiology in young people gives us an opportunity to examine the underlying factors that make people vulnerable to infection.

## 1.3. National HIV Context

The HIV prevalence in Zambia is one of the highest in the region with one in seven adults infected with HIV. The 2007 Zambia Demographic and Health Survey (ZDHS) estimated the adult HIV prevalence as 14.3%. However there are considerable differences by age, gender, region and socioeconomic position (SEP) indicators. A higher proportion of women (16%) are infected with HIV compared to 12% for men. The age specific HIV prevalence pattern differs among women and men. HIV prevalence rises sharply from 6% among women aged 15-19 years to a peak of 26% in women aged 30-34, then declines to 12% in those aged 45-49. Among men, HIV prevalence rises from 4% for men aged 15-19 and peaks at a later age to 24% among those aged 40-44, then declines to 12% in men aged 55-59 years. Rural-urban differences show that HIV prevalence in urban areas is twice as high as in rural areas (20 and 10 percent, respectively). At provincial level Northern and North-Western provinces have the lowest HIV prevalence (7%), while Lusaka has the highest prevalence (21%). Other highly urbanized provinces have prevalence above the national average, i.e. Central (18%), Copperbelt (17%), Western and Southern provinces (15%) [5].

Zambia was among the first countries in Africa to establish a comprehensive surveillance system on HIV among ANC-attendees [6]. This system has been validated in terms of representativeness of ANC-based surveillance both for prevalence estimation and for measurement of trends by using population-based survey data from selected communities [7-9]. Validation of the system was also undertaken through comparisons with a national survey (2001/2 Zambia Demographic and Health Survey) and showed that the two data types match closely with regards to HIV prevalence [10]. This ANC-based surveillance system together with nationally representative population-based surveys (repeated every five years) represent a comprehensive HIV epidemiological data basis to generate knowledge on dynamics and trends of the HIV epidemics in Zambia. There is a realization by the government and partners involved in the implementation of HIV programmes that in order to respond effectively to the epidemic, it is important to "know your epidemic". Emerging issues in HIV prevention include the importance of considering underlying structural issues such as socioeconomic factors at the community and individual level.

## 1.4. Significance Of Socioeconomic Position

There are disparities in the health status of people by socioeconomic position within populations. Systematic differences in health that are avoidable, unjust and unfair are termed 'health inequities'. 'Health equity' is defined as a fair possibility for everyone to access health services and achieve a healthy life [11]. Bravemen and Gruskin in their conceptual paper on equity argue that the term 'avoidable' is redundant, because 'unjust' and 'unfair' implies avoidable. They defined 'equity' as the absence of disparities in health that are systematically associated with social disadvantage or advantage [12]. The recognition of inequalities in health is important for designing effective health programmes.

Socioeconomic position (SEP), sometimes referred to as socioeconomic status, is an individual's or household's economic and social position in relation to others within society based on factors like income, education, occupation and wealth [13]. Most studies investigating the association between SEP and health outcomes show that the better off tend to do better on most measures of health status.

People's structural locations or positions within society are powerful determinants of the likelihood of health damaging exposures or of possessing particular health enhancing resources.

The social science and social epidemiology literature consistently treats Socioeconomic Position (SEP) as a multidimensional construct comprising diverse socioeconomic factors, these different socioeconomic factors could affect health at different times in the life course, operating at different levels (individual, household, neighbourhood) [14]. Socioeconomic position is divided into indicators of economic status and social status. Principal indicators of economic status include individual or household income, household consumption and household wealth. Both income and consumption data are expensive and difficult to collect, particularly in low-income countries, and most health studies lack these direct measures of economic status. Income, which is defined as earnings from productive activities such as labour or services, the sale of goods or property or profit from financial investment, requires obtaining data on all sources of income [15]. In low-income countries a large proportion of the population work in the informal sector, households have multiple and continually changing sources of income and home production is widespread. Therefore surveys have considerable difficulties in collecting accurate data. Compounding this issue is the widespread reluctance to disclose information on income to survey enumerators [15]. In these contexts, consumption expenditure is considered a more stable measure than income. Consumption refers to resources actually used. However, collecting full consumption expenditure data is time consuming. It involves administering lengthy questionnaires on a wide variety of expenditures and it typically takes an hour to complete. In addition, complex calculations and assumptions are required to value home produced goods and account for differences in prices across areas and time for expenditure figures [16].

An alternative proxy measure of household wealth is constructed from data on household assets, housing characteristics and access to services [17]. According to the World Bank the strength of using an asset-based wealth index is that, like consumption data it gives indication of long-term command over resources [15]. One limitation of the wealth index is that it has a residential bias, since assets and services included in the construction of the index are typically owned by urban households but not rural households. One effect of this is that urban households are more likely than rural households to be clustered in the wealthiest quintile. Another concern is that it is difficult to distinguish the poorest of the poor from the other poor, especially in rural areas [18]. Measures of social status often used as indicators of SEP are level of education and employment status.

## 2.0. BACKGROUND TO STUDY OF ASSOCIATION BETWEEN SOCIOECONOMIC POSITION AND HIV

Health research has long been concerned with socioeconomic inequalities in population health status and access to health care services. In epidemiologic research on HIV, the association between socioeconomic factors and HIV transmission has become a critical research component. Studies have focused on indicators of SEP such as education, wealth status, occupation and employment status, but most are limited to individual level SEP.

Gillespie et al, review findings of studies done in Africa between 2004 and 2007, examining the relationship between economic status and the risk of HIV infection. They assess whether epidemics are transitioning from an early phase in which wealth was a primary driver, to one in which poverty is increasingly implicated. This study concluded that the notion that poverty is the main driver of HIV transmission was too simplistic. Relative wealth appeared to have mixed influence on HIV risk depending on an array of contextual factors such as gender inequality, community factors, and access to major roads or transit points. They noted that the role and influence of social cohesion and community-level structural factors is under-researched and little understood, but is very important in investigating the relationship between socioeconomic position and HIV [19].

Mishra et al. examined the association between household wealth status and HIV prevalence in Sub-Saharan Africa using data from nationally representative population-based surveys: 6 Demographic and Health Surveys (Kenya, Malawi, Lesotho, Cameroon, Ghana and Burkina Faso) and two AIDS indicator surveys (Tanzania and Uganda) conducted during the period 2003 – 2005 [20]. Wealth was measured using an asset based index and HIV testing was carried out using dried blood from a finger prick and in one survey venous blood was collected. The overall prevalence in the eight countries ranged from 1.8% in Burkina Faso to 23.5% in Lesotho. In all the countries, HIV prevalence tended to be much higher among adults belonging to the wealthiest 20% of households than among those from the poorest 20%. However, accounting for a number of underlying factors such as education, urban/rural residence and community wealth and for mediating factors using multivariate logistic regression, considerably diminished the positive association between wealth status and HIV in most cases. They concluded that much of the positive association between wealth and HIV was caused by underlying and mediating factors [20].

In a population-based open cohort study on HIV incidence and poverty in Manicaland, Zimbabwe, Lopman et al. found that HIV incidence was lower in the top wealth index tercile (15.4 per 1000 person-years) compared with the lowest tercile (27.4 per 1000 person-years) among men. The pattern of decreasing prevalence trend by wealth was still significant after controlling for site type and age. The pattern was more marked for young men aged 17-24 years among whom incidence was 8.3 per 1000 person years in the highest wealth group and 23.3 in the lowest. No clear significant pattern in incidence by wealth was observed for women of all ages or young women. They concluded that high incidence was associated with poverty in men, especially young men [21].

Barnighausen et al. used data from a cohort study done in a rural community in KwaZulu Natal, South Africa, to investigate the effect of three measures of SES: educational attainment, household wealth and per capita household expenditure, on HIV incidence. This study showed different findings, from the cohort cited above. People belonging to the middle 40% of households as ranked by the asset index had a relative hazard of HIV seroconversion of 2.03 compared to the poorest 40%. Controlling for residence, migration status, partnership status, age and sex reduced the size of the hazard ratio to 1.72, but the effect remained significant. Further, an additional grade of educational attainment reduced the hazard of HIV seroconversion by about 7%. After controlling for sex, age, wealth, household expenditures, place of residence, migration status and partnership status this effect remained significant. However, household expenditure was not a significant determinant of HIV seroconversion. The authors concluded that overall in this poor community it was not the members of the asset poorest households who were at highest risk of HIV infection but people in the middle wealth quintile, and education had a protective effect in relation to the risk of becoming infected with HIV [22].

Education is the most studied socioeconomic variable. In general it appears to be protective with regard to HIV risk as shown in the study by Barnighausen et al. However in the early part of the HIV epidemic, higher educational attainment was associated with a greater risk of infection. Fylkesnes et al. in their study investigating socio-demographic prevalence patterns among childbearing women in Zambia in 1994 found that among both urban and rural residents the seroprevalence was found to be rising significantly with increasing educational attainment [7].

Michelo et al. assessed changes over time in the association between educational attainment and HIV infection in the general population using data from serial population based HIV surveys conducted in selected urban and rural communities in Zambia between 1995 and 2003. The findings revealed a changing pattern of HIV prevalence by educational attainment due to a marked reduction in HIV prevalence among more educated young people, whereas there was stability or even increase among groups with low educational attainment. In urban areas higher educated young men and women had reduced odds of infection than the least educated ((OR 0.20, 95% CI 0.05 – 0.73) and (OR 0.33, 95% CI 0.15 – 0.72), respectively). A similar pattern was observed in rural young men (OR 0.17, 95%CI 0.05 – 0.59) but less prominent and not statistically significant in rural women [23].

Hargreaves et al. review findings of studies done in Africa examining the relationship between educational attainment and HIV infection. The objective of the systematic review was to explore time trends in the association between educational attainment and risk of HIV infection in sub-Saharan Africa. Thirty six articles were included in the study, containing data on 72 discrete populations from 11 countries between 1987 and 2003, representing 200,000 individuals. They found a shift in the association from positive in the early stages to later becoming negative, with higher risk of HIV infections among the least educated. Data on trends in HIV prevalence over time were limited, but findings were consistent with what was observed in the Zambian data, that HIV prevalence fell more consistently among highly educated groups than among less educated [24].

Msisha et al examining the relationship between socioeconomic status and HIV prevalence in Tanzania using data from a national population survey, found that employment status was positively associated with HIV [25]. However, the relationship was substantially different for men and women. Women working in professional jobs had the highest probability of having HIV (OR 1.54, 95% CI 1.17 – 4.82). In contrast, unemployed men were more likely to be HIV positive (OR 3.49, 95% CI 1.43 – 8.58). They suggest that in Tanzania unemployment causes men to travel and migrate, especially from rural to urban areas in search of employment opportunities, which puts them in contact with high risk sexual networks and provides them with the opportunity to engage in casual sexual relationships thereby increasing their likelihood of contracting HIV.

Most studies on SEP and HIV have focused on individual-level factors. Wojcicki in her review of HIV studies in East, Central and Southern Africa found that in 2001 only two out of 36 studies had taken into account neighbourhood socioeconomic factors. She recommended that future studies should examine SEP at both the individual and ecological level [26], but research is still limited. However, interest in the effect of neighbourhood factors on HIV is increasing. Studies examining the issue of context in public health have focused on the magnitude of neighbourhood or area effects on outcomes such as infant and child health, women's health, cardiovascular disease, mortality and health behaviour [27]. For example, Geronimus reported that the effect of maternal aging on birth weight is magnified among African American women who reside in low-income, as compared with high-income, urban areas, suggesting that the impact of maternal age on birth outcomes may depend on some underlying processes associated with social or residential context [28, 29]. Findings such as these have prompted researchers to employ contextualized, multilevel analytic strategies to take into consideration social-structural influences on health [29].

Kayeyi et al. examined the effects of neighbourhood-level educational attainment on HIV prevalence among young women in selected communities in Zambia. This study re-analysed data from a cross-sectional population survey conducted in 2003. A measure for neighbourhood-level educational attainment was constructed by aggregating individual-level years-in-school. Multilevel mixed effects regression models were run to examine the neighbourhood effect on HIV prevalence while adjusting for individual-level underlying variables (age, education, marital status, and currently school attendance). They found that HIV prevalence decreased substantially by increasing level of neighbourhood education. The adjusted likelihood of infection in low vs. high educational attainment of neighbourhoods was 3.4 times (95% CI 0.09 – 0.87) among rural women and 1.8 times (95% CI 0.32 – 1.02) higher among the urban women. After adjusting for the level of education in the neighbourhood, there was a strong protective effect among urban women but not among rural women. They conclude that neighbourhood educational attainment was a strong determinant of HIV infection in both urban and rural young women [30].

Gabrysch et al. in a study aimed at investigating the role of neighbourhood socioeconomic factors on HIV risk in young women in Ndola, Zambia, using population based sero-survey data of nearly 2000 adults conducted in 1997/1998 in a selected community. Neighbourhood-level socioeconomic status (SES) was derived using the availability of running water and electricity in addition to educational, employment and occupational characteristics of adults in the community. Underlying risk factors such as lower and middle SES neighbourhood, having a market nearby, occupation and alcohol use were associated with HIV infection. Young women from middle and lower SES neighbourhoods had 2.4 (95% CI 1.4 – 4.3) and 2.3 (95% CI 1.3 – 4.2) higher odds of HIV infection compared with those from higher SES neighbourhood. Controlling for neighbourhood factors revealed a positive association between higher individual educational level and HIV infection that had not been apparent before. This indicates that ignoring population level factors might lead to different conclusions concerning the role of individual-level factors due to the possibility of confounding effects. The authors concluded that individual-level and population-level factors are quite distinct concepts and can have different effects on HIV infection [31].

## 2.1. RATIONALE

There has been limited research investigating the effect of neighbourhood socioeconomic position on HIV prevalence, and the theoretical frameworks and methodological approaches have varied widely. Information on how underlying community socioeconomic factors are associated with HIV prevalence has important policy implications for prevention programmes and health promotion strategies. Two studies conducted among young women in Zambia found that community level factors were as strongly associated with HIV infection as individual factors [30, 31]. There are, however, few studies that have examined the effect of neighbourhood factors using nationally representative data. This study will re-analyse data from the Zambia Demographic and Health survey using multi-level modelling techniques to examine the relationship between different measures of SEP and HIV prevalence at both individual and neighbourhood level. The study is restricted to young people since most infections in this group can be assumed to have been acquired recently and mortality is low, particularly after the introduction of antiretroviral treatment. The differential prevalence patterns in this age group provide information about recent transmission patterns in the population. Access to ARVs in Zambia was not scaled up before 2003. Thus horizontal transmission can be disregarded among young adults aged 15-24 years in 2007 since children infected with HIV who did not receive ARVs may not survive for that long.

# 3.0. RESEARCH OBJECTIVES

The study will contribute to the body of evidence on the impact of underlying distal socioeconomic factors and HIV prevalence.

## 3.1. MAIN OBJECTIVE

To examine the effect of individual and neighbourhood socioeconomic position factors on HIV prevalence, with a focus on young people aged 15-24 years, in Zambia.

## 3.2. SPECIFIC OBJECTIVES

(1) To examine the effect of educational attainment at individual and neighbourhood level on HIV prevalence in young people aged 15-24 years

(2) To examine the effect of individual and neighbourhood level wealth on HIV prevalence in young people aged 15-24 years.

(3) To examine the effect of employment at individual and neighbourhood level on HIV prevalence in young people aged 15-24 years

(4) To examine the effect of a composite SEP indicator at individual and neighbourhood level on HIV prevalence in young people aged 15-24 years.

# 4.0. STUDY DESIGN AND METHODOLOGY

## 4.1. SETTING

Zambia is a landlocked sub-Saharan country located in Southern Africa with a total land area of 752,612 sq. km. The current population is estimated at 13 million people with an annual growth rate of 2.8. More than a third (39%) of the population live in urban areas. Lusaka and Copperbelt had the highest population, accounting for 40% of the total population. Women account for 51% of the population [32]. About two thirds of the population is below the age of 25 years with 22% being between 15-24 years [33]. The estimated HIV prevalence among those aged 15-49 years is 14.3% and among young people aged 15-24 years it is 6.5% [5].

## 4.2. STUDY DESIGN

This study is based on data from the 2007 Zambia Demographic and Health Survey (2007 ZDHS), a cross-sectional nationally representative population-based study. The sample for the 2007 ZDHS was designed to provide estimates of population and health indicators at the national and provincial levels.

A representative sample of 7,969 households was selected, using a two stage stratified systematic cluster sampling procedure. In the first stage 320 Standard Enumeration Areas (SEAs) were selected from 18 strata with probability proportional to size. Stratification involved separating urban and rural clusters in the 9 provinces. A SEA is a convenient geographical area with an average size of 130 households or 600 people. A complete listing of houses in the selected SEAs was then conducted, and this formed the sampling frame for the selection of households in the second stage. In the second stage, an average number of 25 households were selected in every cluster, by equal probability systematic sampling. The detailed methods of the survey are reported in the 2007 Zambia Demographic and Health Survey report [5].

The survey collected data on socio demographic background variables, fertility, child health, maternal health, sexual behaviour, knowledge, attitudes and practices regarding HIV and STIs, adult mortality and maternal mortality, using structured questionnaires administered by trained interviewers. The instruments used in the survey were pretested in a study of 151 households. The household questionnaire was administered first to the head of household or a knowledgeable adult available at the time of the house visit. Then the individual questionnaire was administered to eligible adults identified from the first interview i.e. women aged 15 to 49 years and men 15-59 years who stayed in the house. The data collection took place over a period of six months, from April to October 2007.

HIV specimens in the form of dried blood spots, were collected from individuals who consented to HIV testing after the individual interview. The protocol for HIV testing allowed for anonymous linking of the HIV results to the socio-demographic data of individuals interviewed. In a sub-sample of one in every three households, syphilis specimens were also collected. Testing was performed among eligible women and men who consented to the test.

## 4.3. Study Population

All women aged 15-49 years and all men aged 15-59 that were either permanent residents of the sampled households or visitors present in the households on the night before the survey were eligible to be interviewed. Each cluster had an average sample of 60 eligible respondents. A total of 7,146 women and 6,500 men were successfully interviewed, yielding response rates of 96.5% and 91.0%, respectively. The principal reason for non-response among eligible adults was failure to find individuals at home despite repeated visits, followed by refusal to be interviewed. Among respondents who completed the structured interview, 75% consented to the HIV testing. Coverage rates for testing were higher for women than for men (77 and 72 percent, respectively). Therefore in total 10,444 respondents 15-49 years were tested. Among these 4253 were aged 15-24 years.

## 4.4. Ethical Considerations

The survey obtained ethical approval from the Tropical Disease and Research Centre (TDRC) ethical committee and the Centre for Disease Control and Prevention (CDC) Atlanta research ethics review board. Participation in the survey was based on informed and voluntary consent, and separate consent was sought for HIV and syphilis testing. Participants were informed that HIV testing was anonymous and was strictly for research purposes. They were also informed about available Voluntary Counseling and Testing (VCT) sites in their area, if they wanted to know their HIV status. Syphilis testing was not anonymous because treatment of seropositive cases was part of the protocol. Participants were informed about this in accordance with ethical requirements.

## 4.5. Operational Definitions

### 4.5.1. Dependent Variable and predictor variables

The dependent variable (HIV) is  a dichotomous variable indicating serostatus determined using blood samples collected from each consenting individual (0 indicating "HIV negative" and 1 "HIV positive"). Major predictor variables are Socioeconomic Position (SEP) indicators represented by three distinct variables in this study: wealth index, educational attainment and employment status. A composite construct of SEP was also created. The variables were defined as:

### 4.5.2 Individual predictor variables

Wealth Index: The ZDHS has a national wealth index as a background characteristic of surveyed respondents. A wealth score was constructed by using principal component analysis based on household assets, housing characteristics and access to amenities data (e.g., roof and floor material, electricity, water supply, possession of goods such as a bicycle and television, and so forth). The general methodology used to calculate the wealth index is based on the Filmer and Pritchett approach [18]. The score was then disaggregated into wealth quintiles (from the poorest to the wealthiest). One limitation of the national quintile approach of relative wealth used in ZDHS is that it has a residential bias, since assets measured are over-represented by assets typically owned by urban households. One effect of this is that urban households are more likely than rural households to be clustered in the wealthiest quintile. Therefore it becomes difficult to distinguish between the poorest of the poor and other poor households, or to make comparisons between relative wealth in rural and urban areas [18]. One possible method to reduce the residence bias is creating separate categories for urban and rural populations, which was done in this study.

Educational attainment: To measure educational attainment, respondents were asked what the highest level of school they attended was and how many years they had spent at that level. The recorded number of years spent in school was used in this analysis as a continuous variable.

Employment status: The variable was measured by asking respondents if they have been working in the past 7 days or in the 12 months preceding the survey. Accurate assessment of employment status can be difficult because some individuals work in the informal sector, family business or are self-employed. To avoid underestimating a respondent's employment status the survey asked several questions including what the respondent mainly did in the past 7 days or in the last 12 months. Depending on the response, the respondent was categorised as employed or unemployed. Categories such as self-employed and unpaid family workers were considered as employed. However housewives/homemakers and students were considered unemployed. Employed individuals were those who had worked at any time during the 12 months preceding the survey [5]. The employment status was categorized as unemployed and employed.

Composite SEP variable: Using the three SEP indicators a composite SEP indicators was constructed. The first step was to categorise the SEP variables. Educational attainment was categorised into four levels according to Zambia's educational system: Primary school (1-7 years), junior secondary school (8-9 years), senior secondary school (10-12 years), and tertiary education (above 12 years). The wealth index was divided into quintiles from one (lowest) to five (highest). The employment status, recoded as 1 if employed and 0 if unemployed. The second step was to sum the three variables. The SEP composite indicator is therefore a linear combination of wealth index, educational attainment and employment status. The resulting composite SEP variable was grouped into three similarly sized groups.

### 4.5.3 COVARIATES

Age, sex, marital status and residence were considered potential confounders or effect modifiers.

### 4.5.4. NEIGHBOURHOOD PREDICATOR VARIABLES

In this study neighbourhoods are defined as geographic areas with similar characteristics, located in close proximity. A cluster (enumeration area - defined as the smallest geographic unit that demarcates a country and used in the census of populations and other types of surveys) was used as a proxy for a neighbourhood.

The neighbourhood variables were derived by aggregating (calculating mean values for each cluster) the following socioeconomic variables: wealth index, education and employment status of all the respondents aged 15-59 years. The neighbourhood variables were constructed based on an aggregation method in STATA 11.

Neighbourhood Wealth**:** The relative wealth status for each neighbourhood was derived by calculating the mean of the wealth scores of all respondents aged 15-59 in the clusters. The clusters were grouped into three categories (low, medium and high relative wealth). The categorization was done separately for urban and rural areas and the group cut-off points were such that the groups were of similar sizes.

Neighbourhood Education**:** The neighbourhood-level educational attainment is estimated by calculating the mean number of years in school for all respondents aged 15-59. The clusters were grouped into three categories of similar size; low, medium and high neighbourhood educational level. Rural-urban differences were taken into account by creating different categories for urban and rural areas.

Neighbourhood Employment Status: The neighbourhood employment status is derived by calculating the proportion of employed respondents aged 15-59 for each neighbourhood. The clusters were categorized into three groups of similar size; low, medium and high neighbourhood employment status. The categorization was done separately for urban and rural areas.

Neighbourhood composite SEP variable: A neighbourhood composite index of socioeconomic position was derived by aggregating the individual composite index in the clusters. The clusters were categorized into three groups of similar size; low, medium and high composite SEP. The categorization was done separately for urban and rural areas.

# 5.0. DATA ANALYSIS AND ESTIMATION PROCEDURES

## 5.1. CONCEPTUAL FRAMEWORK

The analysis and interpretation of data was based on the proximate determinants framework for factors affecting the risk of transmission of HIV developed by Boerma and Weir. The critical aspects of the framework are that underlying variables, such as socioeconomic status, influence proximate determinants, which in turn have a direct effect on biological mechanisms to influence health outcomes (i.e. acquisition of HIV infection). The underlying variables in this study include the socioeconomic context of the neighbourhood (neighbourhood educational attainment, wealth index and employment status) and individual factors (wealth index, educational attainment, employment, age, sex, marital status and residence). The proximate determinants include among others age at first sex, number of sexual partners and non-regular sexual partnerships. The conceptual framework helps to understand the causal pathways from distal socioeconomic factors to HIV infection. It is useful if HIV surveys are guided by this conceptual framework for the purpose of generating knowledge about the epidemiological context of HIV and to be able to determine the suitability of different preventive interventions[34].

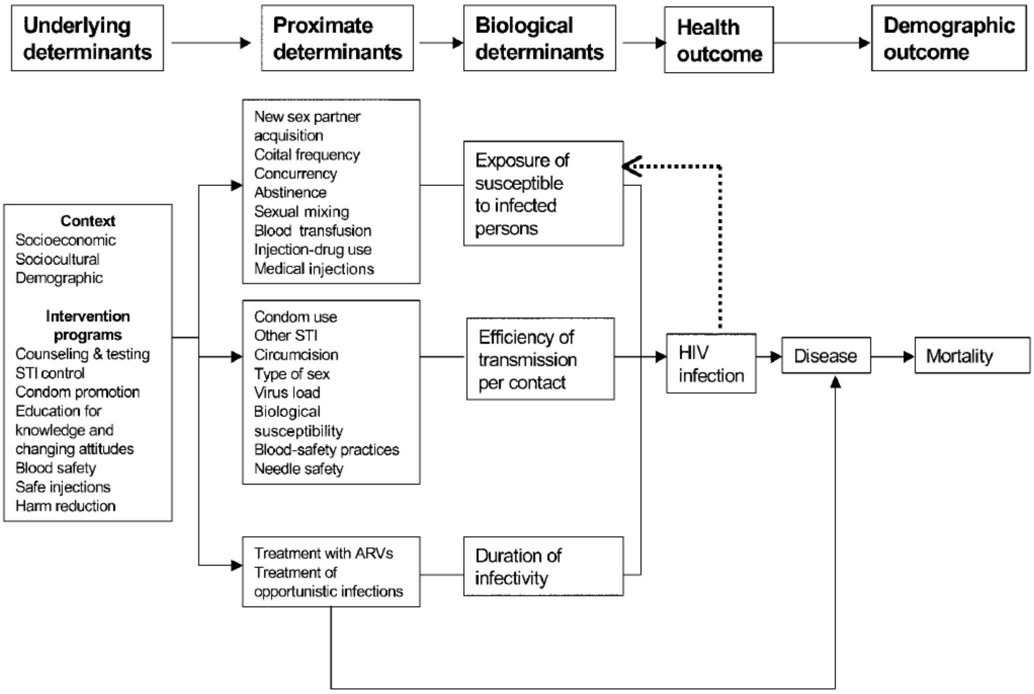



**Figure 1.** Proximate-determinants conceptual framework for factors affecting the risk of sexual transmission of HIV. ARVs, antiretrovirals; STI, sexually transmitted infection.

Using the conceptual framework we systematically examined the association between each SEP Indicator and HIV infection. In our analysis we also took into account that other underlying background factors influence the association between SEP indicators and HIV infection as conceptualised by the framework. These underlying factors include community level factors and demographic factors (Sex, age, marital status). Boerma and Weir suggest that derived group level variables have shown to be important underlying structural factors for HIV infection [34]. In our analysis, neighbourhood SEP indicators are used as proxies for underlying socioeconomic level of the community.

## 5.2. STATISTICAL ANALYSIS

The analysis was restricted to respondents with known HIV status and was stratified by residence. The characteristics of the study population in terms of demographic and socio–economic variables were analysed using descriptive statistics. Analysis was conducted using STATA version 11. All the descriptive analysis and HIV distribution were adjusted for complex study design using the svy command in STATA.

The association between the socioeconomic factors and HIV prevalence was investigated using bivariate and multivariate mixed effects logistic regression models. Variables included in the regression model were specified 'a prior' based on literature review and guided by the proximate determinants framework. The bivariate and multivariate analysis also controlled for the potential confounder age, which was adjusted for as a linear effect. Multilevel mixed- effect logistic regression models estimated both fixed and random effects and took into account clustering of the data. The formula is given by

$$y_{ij} = \begin{cases} \beta_0 + \beta_1 x_{ij} \\ fixed\ part \end{cases} + \begin{cases} u_j + e_{ij} \\ random\ part \end{cases} \qquad \text{Eqn. (i)}$$

*Fixed Part*:

Parameters estimated were coefficients similar to ordinary regression equation

$\beta_0$ - Intercept

$\beta_1$ - Coefficient of the explanatory variable

*Random Part*:

Parameters estimated were Variances

$u_j$ - Variance at cluster level

$e_{ij}$ - Variance at individual level

Most standard statistical tests are based on the assumption of independence of observations. Violation of this assumption affects estimates of standard errors. However social research has shown that individuals are not just operating independently but are influenced by the social groups or context to which they belong [35]. A multilevel model takes account of the hierarchical structure of the survey data structure.

To assess the whether individual risk factors were likely to be mediated by neighbourhood risk factors and vice versa, three separate models were constructed. Model 1 only included individual socioeconomic predictor variables. Model 2 included both neighbourhood and individual SEP variables. The final model was a full model that included all variables and covariates (i.e. sex, marital status) in the same model. Further, using the likelihood-ratio test, we tested for a statistically significant difference between model 1 and model 2, to assess whether adding neighbourhood variables improved the goodness of the fit. Separate models were constructed for urban and rural neighbourhoods.

The study also examined the association between underlying socioeconomic position and HIV infection using three approaches. We used three approaches in order to examine different options of handling socioeconomic indicators and their association with on HIV prevalence. SEP indicators are often handled separately because of concerns of colinearity and are assumed to be inter-changeable proxies of SEP. Hence, some studies have used only one SEP measure as a proxy for socioeconomic position at individual and /or neighbourhood level or used a SEP construct. However, Braveman et al. in a review paper socioeconomic status in health research conclude that different SEP measures cannot be assumed to be interchangeable. Furthermore, a composite SEP does not permit study of how particular socioeconomic factors influence health outcomes [14]. The first approach was restricted to educational attainment at individual and neighbourhood level. The second approach examined three socioeconomic indicators, education, wealth and employment status in multivariate models at both individual and neighbourhood level. The third approach examined a composite SEP indicator constructed using education, wealth status and employment levels both at individual and neighbourhood level.

# 6.0. RESULTS

## 6.1. General Characteristics of Study Population

The distribution of participants used in our current analysis is shown in Table 1. In total 4,253 young people were interviewed and consented to the HIV testing. The percentage living rural areas was 52 percent, while 48 percent were residing in urban areas. There were more young women than men in the study population in both rural and urban areas. Marital status differed markedly by residence with a higher proportion of respondents in rural areas being married (34%) compared to urban areas (17%). Only 3% of young people reported to be formerly married (i.e. separated, divorced), and there were no differences by residence. There was also a striking difference in educational attainment by residence. More than two thirds (72.5%) of urban residents had secondary or higher education, while less than a third (30.1%) in rural areas had attained the same level of education. A substantial proportion in rural areas had only attained primary education (63.3 %). Similar to educational attainment, employment status differed by residence. More than half (55%) of young people in rural areas were employed, compared to slightly over a third (34.4%) in urban areas. In terms of wealth, there were more young people in the wealthiest category in urban than rural areas (44% compared to 34%) (Table 1).

An estimated 6.5% of young people tested HIV positive. The prevalence was higher in urban than rural areas (8.5% compared to 4.7%). A higher proportion of young women (8.5%) were infected with

HIV compared to men (4.3%). The sex differentials by residence show that HIV prevalence was much higher for young women living in urban areas (11.2%) compared to rural areas (6.2%). A similar pattern was observed for men (5.7 and 2.9%, respectively). HIV prevalence was higher among respondents with no education (7.6%) compared with those with primary education (6.2%). There was a marked difference in HIV prevalence by SEP indicators, particularly by residence. In urban areas prevalence declined with an increase in individual educational attainment. About one in five (19%) young people with no education were HIV positive, while those with secondary and higher education had the lowest prevalence (7.1%). In rural area, it was those with primary education that had the lowest prevalence. Individuals with middle wealth status in urban areas had the highest prevalence (10.9%) and the lowest prevalence was observed in the wealthiest individuals (6.2%). However in rural areas the wealthiest had the highest prevalence (6.5%). Being employed in rural areas was associated with low prevalence (3.7%), and the unemployed had a prevalence of 5.9%. In contrast the unemployed had the lower prevalence in urban areas (7.5%), while the employed had a prevalence of 10.3% (Table1).

## 6.2. Factors associated with HIV Prevalence

Results from the bivariate and multivariate mixed effects logistic regression show that education was associated with HIV prevalence but the effects differed by residence (Table 2). Neighbourhood educational attainment tended to be associated with reduced odds of infection (Figure 2), but the association was not statistically significant. In contrast, in rural areas young people living in neighbourhoods with medium compared to low educational attainment that had the highest odds of being infected (OR 2.88, 95% CI 1.54 – 5.38)). At individual level, on average, an additional year of educational attainment reduced the odds of being infected in urban areas (OR 0.93, 95% CI 0.88 – 0.98). In rural areas education tended to be associated with increased odds of HIV infection (OR 1.06, 95% CI 0.98 – 1.13). Adjusting for level of education in the neighbourhood did not affect the protective effect of individual education in urban areas, while in rural areas the association was weakened further (OR 1.03, 95% CI 0.95 – 1.11). In this approach, which was restricted to education as a SEP variable, inclusion of neighbourhood level education in model 2 significantly improved the fit in rural but not urban areas. Including marital status and sex in the final model weakened the protective effect of education on HIV infection in urban areas. Apart from some evidence of statistically significant interactions between education and sex in rural areas, there were no significant interactions between education and other factors.

**Figure 2: Age adjusted odds ratio of HIV infection by neighbourhood educational attainment**

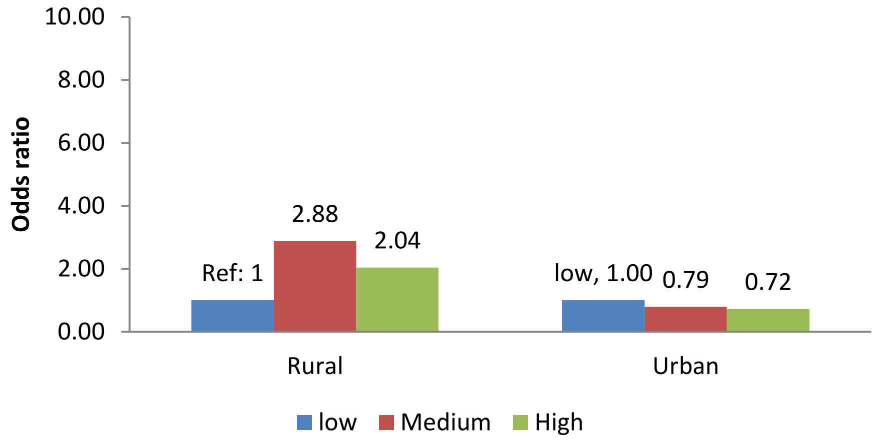
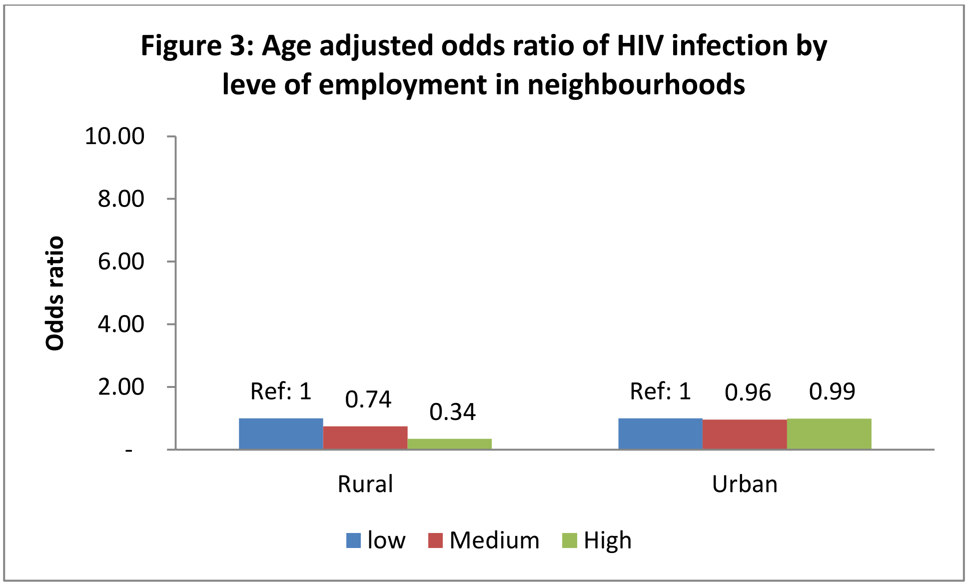


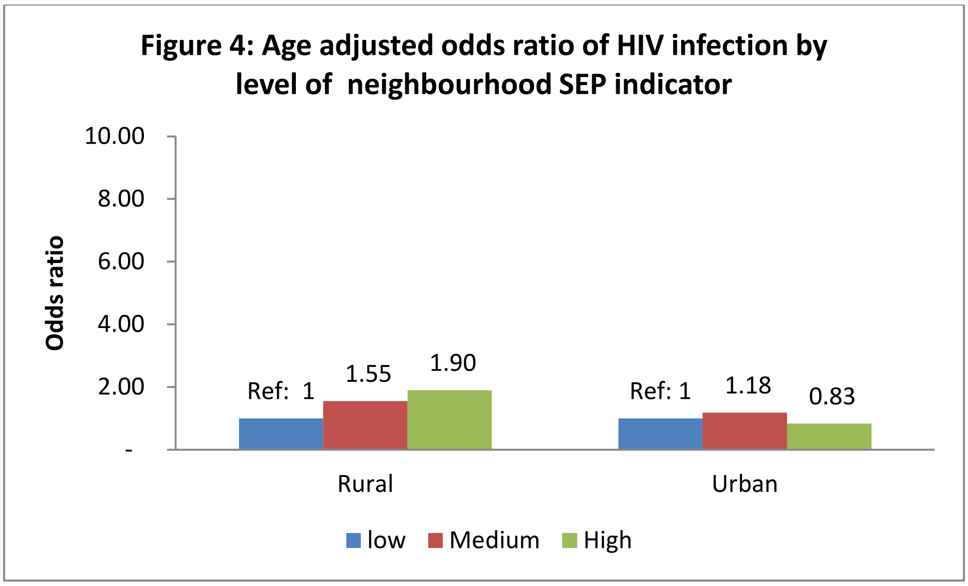

Table 3 presents data on bivariate and multivariate analysis (based on multilevel mixed effects logistic regression models) of all SEP indicators and HIV prevalence. Educational attainment and employment were associated with HIV prevalence. Neighbourhood employment was associated with reduced odds of HIV infection in rural areas (Figure 3), i.e. living in a neighbourhood with high average employment was protective for young people (OR 0.38, 95% CI 0.20 – 0.73). After adjusting for individual and other neighbourhood SEP indicators, the association remained significant. No statistically significant association was observed in urban areas. Employment status of the individual was a strong factor of HIV infection in both rural and urban areas. Young employed people had reduced odds of being infected (OR 0.53, 95% CI 0.34 – 0.81) in rural areas and (OR 0.65, 95% CI 0.46 – 0.91) in urban areas. Educational attainment was also protective for young people living in urban areas. Adjusting for SEP neighbourhood variables weakened the protective effect of individual employment in rural but not urban areas. There was little change in the protective association of individual education and HIV infection in urban areas, and it remained significant.




High average neighbourhood wealth was associated with increased odds of infection in rural areas. However, after adjusting for individual wealth and other SEP indicators the association with HIV prevalence was not statistically significant. In this second approach, which examined all SEP indicators in the regression models, inclusion of neighbourhood level variables in model 2 significantly improved the fit in rural areas, but not urban areas. Including marital status and sex in the final model reduced the effect of all individual SEP indicators on HIV infection in rural and urban areas (Table 3). However, the protective effect of high average neighbourhood employment remained significant in rural areas. There were no statistically significant interactions between SEP variables and other risk factors, apart from some interaction between employment and marital status.

Furthermore we assessed for cross-level interactions between neighbourhood SEP and individual SEP variables but found no statistically significant interactions. Consistent with other studies, education and wealth had a strong correlation of 0.52, and there was a negative but weak correlation between employment and education. We however checked and found no evidence of multi-colinearity in both urban and rural areas.

Higher socioeconomic position of neighbourhoods measured by the composite socioeconomic indicator was positively associated with HIV prevalence in rural areas (Table 4 and figure 4). Individuals in rural areas residing in high compared to low average socioeconomic neighbourhoods were more likely to be infected (OR 1.90, 95% CI 1.05 - 3.45), whereas living in neighbourhoods of high SEP tended to be protective in urban areas, but the association was not significant. Adjusting for individual SEP weakened the significant association of neighbourhood SEP and HIV in rural areas. However, in urban areas residing in a high SEP neighbourhoods changed from being protective to being a risk factor, but the association was not significant. Young people residing in urban areas with high SEP had reduced odds of infection (OR 0.50, 95% CI 0.30 – 0.82). This protective effect of individual SEP on HIV risk remained significant even after controlling for neighbourhood SEP. In contrast, in rural areas high SEP was not significantly associated with HIV prevalence. In the final model the protective effect of individual SEP on HIV infection reduced after adding marital status and sex, while in rural areas the effect remained statistically insignificant. There were no statistically significant interactions between SEP and other risk factors.




## 7.0. DISCUSSION

The findings from this nationally representative population-based survey indicates that there are differences in the likelihood of HIV infection associated with socioeconomic position at both individual and neighbourhood level in rural and urban areas of Zambia. In rural areas, young people from neighbourhoods with high employment levels were less likely to be infected with HIV. In urban areas, neighbourhood educational attainment tended to be associated with reduced risk of HIV infection, but the association was not significant. At individual level being employed significantly reduced the odds of HIV infection in both rural and urban areas. Educational attainment and SEP were also associated with reduced risk of HIV infection but limited to young people living in urban areas.

To examine the independent and joint effects of individual and neighbourhood SEP factors on HIV infection, we assessed how conclusions vary depending on whether a single measure (education in this case), all SEP variables or a composite SEP indicator were used. Educational attainment is an important indicator of socioeconomic position and is the most studied SEP variable in studies on health status. We found that individual educational attainment was protective against HIV infection in urban areas, regardless of whether education alone was used or employment and wealth were included as measures of SEP. An additional year of educational attainment reduced the odds of being infected by 7%. Other studies have documented similar findings [22-24]. A systematic review assessing the association between education and HIV found that studies in sub-Saharan Africa done after the epidemic matured showed lower risk of infection among the most educated [24].We also found that in multilevel analysis, after adjusting for underlying neighbourhood educational attainment, the protective effect of individual education in urban areas was maintained. In rural areas, there was no statistically significant association between education and HIV infection. Our findings of the protective effect of education in urban areas are consistent with a study done in selected communities in Zambia in 2003 [30].

In contrast, the association of neighbourhood educational attainment and HIV infection was weakened after adjusting for individual educational attainment in both rural and urban areas (where higher neighbourhood educational attainment tended to be protective) and rural areas (where higher neighbourhood educational attainment was a risk factor), and the adjusted association ended up not being significant for the most educated neighbourhoods. This indicates that the association between neighbourhood educational attainment and HIV status was partially mediated through the individual educational attainment as conceptualised in the proximate determinants framework. Similar findings were observed in a study done in selected communities in Zambia [30]. In rural areas young people in middle education neighbourhoods had the highest risk of infection. This suggests that neighbourhood educational attainment is still a risk factor, but findings suggest that changes towards less risk taking in the most educated neighbourhoods might occur earlier than in the middle neighbourhoods.

The finding that neighbourhood and individual educational attainment had contrasting associations with HIV prevalence in urban versus rural areas may reflect lifestyle and quality of education differences between rural and urban areas. The more educated in rural areas may be more likely to engage in particular risky lifestyles, such as higher rates of partner change and multiple sexual partners, because of greater autonomy and spatial mobility [19, 20]. In rural areas it is also plausible that mobility still plays a role in the educated being more at risk of infection then the less educated. Young people migrate from rural areas to attend secondary and higher education, and this exposes them to a higher HIV prevalence if they engage in sexual activity (since HIV prevalence is higher in urban than rural areas), and thus they are exposed to a higher risk of transmission than those who remain in the village. There is convincing empirical evidence of the link between human mobility and HIV spread. In Sub-Saharan Africa the risk of HIV infection has been found to be higher near roads and amongst those who migrate or have sexual partners who migrate[19] .

In urban areas, one plausible explanation for the protective effect of education is that the most educated tend to be early adopters of new practices, which then diffuses to the rest of the social system. The more educated seem to be responding to the preventive strategies, and thus changing their behaviour. This pattern of behavioural change is part of the proposed process in the Diffusion of Innovation Theory. It states that, the first individuals to adopt an innovation (innovators) usually have the highest social class, youngest in age and are closest to scientific knowledge [36]. Another crucial concept, the 'critical mass', proposes that when enough individuals adopt an innovation, it becomes self-sustaining, hence neighbourhoods with high educational attainment would be more likely to sustain health promoting behaviours. High neighbourhood educational attainment might also be protective because people have better/comprehensive knowledge about HIV prevention and risk. Our findings also support the 'education vaccine' view, which suggests that education is an important factor in reducing new infections [24, 37]. We could have expected similar changes to occur in rural areas but this does not seem to be the case.

In this study including all three SEP variables (wealth, education and employment) in the second approach, enabled us to assess how each particular SEP variable influenced HIV infection, while controlling for other SEP indicators. Neighbourhood employment was a strong protective factor for HIV infection in rural areas. In contrast the study done in selected communities in Zambia showed no significant association between neighbourhood employment and HIV prevalence [30]. The difference may be due to the operational definitions of employment on the one hand. On the other hand there could be a stronger association in our nationally representative population study than what was seen in the selected communities. However, overall the direction of association is the same. At individual level, employment among young people was protective regardless of residence. In multilevel analysis the magnitude of effects was reduced in rural but not in urban areas after controlling for neighbourhood SEP variables, but the associations remain significant. These findings suggest the importance of joint individual and structural effects on HIV prevalence, especially in rural areas. The protective effect of employment may reflect the importance of empowering young people and further strengthen the argument that the more empowered tend to adopt protective practices earlier such as being less likely to engage in risky behaviour.

In addition, in rural areas, the dominant occupation was agriculture with 40% of the respondents working in this sector. Respondents in this type of labour are less likely to be mobile and hence less susceptible to HIV infection (because they are not exposed to the higher prevalence in towns). Table 1 shows that the prevalence was low among those with agriculture occupation (3 %) and high among rural professionals (10%) who are more likely to be mobile. Studies have shown that mobile individuals seem to be more vulnerable to infection [37, 38]. Furthermore, the main agriculture activity is mostly subsistence or small scale farming. Some respondent even reported being unpaid family workers. Subsistence farming provides just enough income for daily sustenance but not extra resources to spend on activities that may put individuals at risk, such as having multiple partners or going to bars, taverns or local drinking places where people may meet new partners [39]. Material exchange has been documented as a component of sexual relationships in many African countries, especially casual and multiple partnerships which put young people at risk of HIV infection [19, 40]. The contrasting effect of the employment and education variables at an individual level in rural areas may in addition be explained by the fact that students are considered to be ‘unemployed’ in the DHS.

Similar to education, neighbourhood wealth was also associated with increased risk for those in the wealthiest group in rural areas, but tended to be protective in urban areas (although not significant). Adjusting for individual and other SEP indicators weakened the effect in rural areas and changed high neighbourhood wealth to a risk factor in urban areas (but the association was still non-significant). There was a strong correlation between neighbourhood wealth and education, suggesting that the two SEP indicators capture partially similar underlying contextual aspects of the neighbourhood. This partly explains the weakened association between neighbourhood wealth and HIV prevalence after adjusting for level of educational attainment in the neighbourhood in rural areas. In urban areas, the change of direction after the adjustment shows that when differences in education between neighbourhoods are accounted for, the wealthiest neighbourhoods are relatively risky for young people. This may mean that if economic resources are more readily available, young people are more likely to engage in risky practices (multiple partnerships, causal sexual encounters, participation in social and sexual networks), and this effect was masked by the protective effect of educational attainment in the neighbourhood (due to the high correlation between education and wealth) [20, 21, 41].

Finally, in our last approach we found that living in a high SEP neighbourhoods (using a composite index) was also associated with increased odds of infection in rural areas. In urban areas the neighbourhoods with high composite SEP had the lowest prevalence compared to low average composite SEP neighbourhoods, but the association was not significant. This finding is not surprising since education and wealth were weighted more on the SEP construct. This is consistent with a study done in Ndola, Zambia, where the authors found significantly reduced risk of infection in high SES neighbourhoods and established that girls in lower and middle SEP neighbourhoods were more likely to have started sexual activity earlier, be married and infected with HSV2 or syphilis. Differences in how the composite SEP was constructed, may account for the observed significant association found in the aforementioned study but not our study. The authors included electricity and water availability in the SEP construct [31]. The difference may also be due to changes in the association over time.

Inclusion of other covariates sex and marital status in the multilevel models reduced the association between neighbourhood and individual SEP and HIV prevalence. Sex is likely to be confounding the association between SEP indicators and HIV prevalence. Young women were less likely to spend more years in school or to be employed than young men. At the same time our study shows that young women were about twice as likely to be infected as men. This finding is consistent with other studies done in Southern Africa that show that HIV disproportionately affects women than men especially in the younger age groups.

There are several limitations to our study. As in any cross-sectional study the exposure and outcome are measured simultaneously in this given population. We are therefore unable to distinguish between the effect of SEP on HIV infection and the effect of HIV infection on SEP. However, since the study is limited to young people, HIV infection is likely to be recent. Some important factors may have biased the neighbourhood effects in our study, such as using a cluster as a proxy of neighbourhoods. Clusters which are defined as census enumeration areas are not necessarily representative of naturally occurring neighbourhoods where individuals reside. Studies have, however, found them to be small enough to be a useful proxy [27, 30]. Furthermore, there was no information on more specific community level characteristics that are associated with HIV prevalence in other studies, such as availability of markets, health facilities, bars, and proximity to trading areas and major roads. Non-participation of 25% is likely to some extent have biased our HIV prevalence estimates. Population based surveys are vulnerable to selection bias if HIV status is missing for a proportion of the eligible population as is the case with our data [42]. However, a study that assessed declines in prevalence among young women in Zambia found that the point prevalence estimates only increased by 1-2 percentage points in the most extreme scenario after conducting a sensitivity analysis [43] . Another study that analysed five population-based surveys found that non-response did not bias national estimates from population based studies significantly, especially in countries with relatively high prevalence. They found that although prevalence is predicted to be higher in men and women who were not tested than those tested in all five countries, the overall effects tended to be small when estimates were adjusted for non-response [44]. In general, minimizing non-response is a major challenge for all population based surveys, particularly surveys that include collection of blood samples for testing. We did not attempt to make adjustments to the prevalence estimates.

Our methodological approach was able to capture different dimensions of SEP and assess its association with HIV prevalence. Further the DHS programme has become a major source of data on HIV prevalence. The DHS surveys been validated in many countries and have stringent data quality assurance processes. The consistent high quality of DHS data provides reliable estimates of HIV prevalence, and is useful for identifying factors that make young people vulnerable to infection.

Neighbourhood socioeconomic factors play a role in young people's vulnerability to HIV infection regardless of whether SEP is restricted to one measure or more SEP structural factors are measured. In the absence of incidence data, studying HIV prevalence in young people shows some of the complexities of the association between SEP and HIV. Our findings also indicate need for further research in order to better capture these complexities and thus provide better preventive guidance.

Different SEP indicators at individual and neighbourhood level had contrasting effects on HIV prevalence in rural and urban areas. Our results also suggest that community level factors have a more important influence in rural than urban areas.

Recommendations

1. Findings from this study could inform preventive strategies. Programme and policy makers need to focus more on preventive strategies to target community level factors in addition to individual prevention efforts.
2. IEC materials should be designed and targeted specifically to communities to compliment already existing IEC materials. Materials must also be audience specific, taking into account that communities or neighbourhoods differ in rural and urban areas
3. Findings in the survey can provide evidence for future HIV policy. For instance moving beyond assumptions such as poverty is a strong driver in HIV transmission as outlined in the current HIV policy in Zambia that led to broad based strategies of mainstreaming HIV in the poverty reduction strategy (PRSP). Although poverty reduction is an essential strategy, our analysis suggests the need for being more specific in addressing structural drivers. An example is the IMAGE Programme in South Africa that combines micro finance for women and community mobilisation to reduce HIV risk behaviour among young women [1].
4. The study highlights the importance of providing and improving education in rural areas. It should be a broad based strategy being complimented with other programmes that empower young people.
5. Further research needs to be conducted to examine the observed contrasting association between education, wealth, employment and HIV prevalence among rural and urban respondents.

## References


1. UNAIDS W: **UNAIDS Report on the global AIDS epidemic 2010**. UNAIDS Geneva; 2010.
2. Central Statistical Office (CSO) MoHM, University of Zambia and MEASURE Evaluation: **Zambia Sexual Bahaviour Survey 2009**. Lusaka: Central Statistical Office and MEASURE Evaluation; 2010.
3. Rodríguez MP, Hayes R: **Reducing HIV prevalence among young people: A review of the UNGASS prevalence goal and how it should be monitored**. *London School of Hygiene and Tropical Medicine, London, UK* 2002.
4. Unicef, Unaids, Unfpa, Ilo, Who: *Opportunity in Crisis: Preventing HIV from Early Adolescence to Early Adulthood*: United Nations Children's Fund, The (UNICEF); 2011.
5. Central Statistics Office (CSO) MoHM, University of Zambia, Tropical Disease Research Centre (TDRC), Macro International Inc: **Zambia Demographic and Health Survey 2007**. Central Statistics Office and Macro International Inc; 2009.
6. Fylkesnes K, Musonda RM, Kasumba K, Ndhlovu Z, Mluanda F, Kaetano L, Chipaila CC: **The HIV epidemic in Zambia: socio-demographic prevalence patterns and indications of trends among childbearing women**. *AIDS* 1997, **11**(3):339.
7. Fylkesnes K, Ndhlovu Z, Kasumba K, Musonda RM, Sichone M: **Studying dynamics of the HIV epidemic: population-based data compared with sentinel surveillance in Zambia**. *AIDS* 1998, **12**(10):1227.
8. Fylkesnes K, Musonda RM, Sichone M, Ndhlovu Z, Tembo F, Monze M: **Declining HIV prevalence and risk behaviours in Zambia: evidence from surveillance and population-based surveys**. *AIDS* 2001, **15**(7):907.
9. Michelo C, Sandøy I, Dzekedzeke K, Siziya S, Fylkesnes K: **Steep HIV prevalence declines among young people in selected Zambian communities: population-based observations (1995–2003)**. *BMC Public Health* 2006, **6**(1):279.
10. Dzekedzeke K, Fylkesnes K: **Reducing uncertainties in global HIV prevalence estimates: the case of Zambia**. *BMC Public Health* 2006, **6**(1):83.
11. Lindstrand A BS, Rosling H, Rubenson B, Stenson B, Tylleskar T: *GLOBAL HEALTH - An introductory textbook*. Edition 1:5 edition: Studentlitteratur; 2006.
12. Braveman P, Gruskin S: **Defining equity in health**. *Journal of epidemiology and community health* 2003, **57**(4):254.
13. Lynch J, Kaplan G: **Socioeconomic position**. *Social epidemiology* 2000:13-35.
14. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, Posner S: **Socioeconomic status in health research: one size does not fit all**. *Jama* 2005, **294**(22):2879.
15. O'Donnell O. DEV, Wagstaff A.,Lindelow M.: **Analyzing Health Equity Using Household Survey Data**. Washington DC.: The World Bank; 2008.
16. Howe LD, Hargreaves JR, Gabrysch S, Huttly SRA: **Is the wealth index a proxy for consumption expenditure? A systematic review**. *British Medical Journal* 2009.
17. Rutstein SO, Johnson K, Macro O: *The DHS wealth index*: ORC Macro, MEASURE DHS; 2004.
18. Rutstein S: **The DHS wealth index: approaches for rural and urban areas**. *Washington, DC: Macro International Inc* 2008.
19. Gillespie S, Kadiyala S, Greener R: **Is poverty or wealth driving HIV transmission?** *AIDS* 2007, **21**:S5.
20. Mishra V, Assche SBV, Greener R, Vaessen M, Hong R, Ghys PD, Boerma J, Van Assche A, Khan S, Rutstein S: **HIV infection does not disproportionately affect the poorer in sub-Saharan Africa**. *AIDS* 2007, **21**:S17.
21. Lopman B, Lewis J, Nyamukapa C, Mushati P, Chandiwana S, Gregson S: **HIV incidence and poverty in Manicaland, Zimbabwe: is HIV becoming a disease of the poor?** *AIDS (London, England)* 2007, **21**(Suppl 7):S57.
22. Barnighausen T, Hosegood V, Timaeus IM, Newell ML: **The socioeconomic determinants of HIV incidence: evidence from a longitudinal, population-based study in rural South Africa**. *AIDS* 2007, **21 Suppl 7**:S29-38.
23. Michelo C, Sandøy IF, Fylkesnes K: **Marked HIV prevalence declines in higher educated young people: evidence from population-based surveys (1995-2003) in Zambia**. *AIDS* 2006, **20**(7):1031.

24. Hargreaves JR, Bonell CP, Boler T, Boccia D, Birdthistle I, Fletcher A, Pronyk PM, Glynn JR: **Systematic review exploring time trends in the association between educational attainment and risk of HIV infection in sub-Saharan Africa**. *AIDS* 2008, **22**(3):403-414.
25. Msisha WM, Kapiga SH, Earls F, Subramanian SV: **Socioeconomic status and HIV seroprevalence in Tanzania: a counterintuitive relationship**. *Int J Epidemiol* 2008, **37**(6):1297-1303.
26. WOJCICKI J: **Socioeconomic status as a risk factor for HIV infection in women in East, Central and Southern Africa: a systematic review**. *Journal of Biosocial Science* 2005, **37**(01):1-36.
27. Roux D, Ana V: **Investigating neighborhood and area effects on health**. *American journal of public health* 2001, **91**(11):1783.
28. Geronimus AT: **Black/white differences in the relationship of maternal age to birthweight: A population-based test of the weathering hypothesis**. *Social Science & Medicine* 1996, **42**(4):589-597.
29. Rauh VA, Andrews HF, Garfinkel RS: **The contribution of maternal age to racial disparities in birthweight: a multilevel perspective**. *American journal of public health* 2001, **91**(11):1815.
30. Kayeyi N, Sandøy IF, Fylkesnes K: **Effects of neighbourhood-level educational attainment on HIV prevalence among young women in Zambia**. *BMC Public Health* 2009, **9**(1):310.
31. Gabrysch S, Edwards T, Glynn JR: **The role of context: neighbourhood characteristics strongly influence HIV risk in young women in Ndola, Zambia**. *Tropical Medicine & International Health* 2008, **13**(2):162-170.
32. (CSO) CSO: **Zambia 2010 Census of Population and Housing Preliminary Report**. Lusaka: Central Statistical Office; 2011.
33. (CSO) CSO: **Zambia 2000 Census of Population and Housing** *Volume 10*. Central Statistical Office, Ministry of Finance and National Planning; 2003.
34. Boerma J, Weir SS: **Integrating demographic and epidemiological approaches to research on HIV/AIDS: the proximate-determinants framework**. *Journal of Infectious Diseases* 2005, **191**(Supplement 1):S61.
35. Hox JJ: *Multilevel analysis: Techniques and applications*: Taylor & Francis; 2010.
36. Rogers EM: *Diffusion of innovations*: Free Pr; 1995.
37. Vandemoortele J, Delamonica E: **The "education vaccine" against HIV**. *Current issues in comparative education* 2000, **3**(1):6-13.
38. Buvé A, Bishikwabo-Nsarhaza K, Mutangadura G: **The spread and effect of HIV-1 infection in sub-Saharan Africa**. *The Lancet* 2002, **359**(9322):2011-2017.
39. Ndubani P, Kamwanga J, Tate J, Buckner B: **PLACE in Zambia: Identifying gaps in HIV prevention in Kapiri Mposhi, Central Province, 2005**. *Chapel Hill, NC: University of North Carolina at Chapel Hill* 2006.
40. Maganja R, Maman S, Groves A, Mbwambo J: **Skinning the goat and pulling the load: transactional sex among youth in Dar es Salaam, Tanzania**. *AIDS care* 2007, **19**(8):974-981.
41. Parkhurst JO: **Understanding the correlations between wealth, poverty and human immunodeficiency virus infection in African countries**. *Bulletin of the World Health Organization* 2010, **88**:519-526.
42. Bärnighausen T, Bor J, Wandira-Kazibwe S, Canning D: **Correcting HIV prevalence estimates for survey nonparticipation using Heckman-type selection models**. *Epidemiology* 2011, **22**(1):27.
43. Kayeyi N, Fylkesnes K, Michelo C, Makasa M, Sandøy I: **Decline in HIV Prevalence among Young Women in Zambia: National-Level Estimates of Trends Mask Geographical and Socio-Demographic Differences**. *PloS one* 2012, **7**(4):e33652.
44. Mishra V, Vaessen M, Boerma JT, Arnold F, Way A, Barrere B, Cross A, Hong R, Sangha J: **HIV testing in national population-based surveys: experience from the Demographic and Health Surveys**. *Bull World Health Organ* 2006, **84**(7):537-545.

# APPENDICES

**Table 1: Percent distribution of young people15-24 years by background characteristics and HIV prevalence, in Zambia**

| | Total | | | Rural | | | Urban | | |
|---|---|---|---|---|---|---|---|---|---|
| **Predictor Variables** | **Number** | **%** | **HIV%** | **Number** | **%** | **HIV%** | **Number** | **%** | **HIV%** |
| Sex | | | | | | | | | |
| Male | 2028 | 47.7 | 4.3 | 1038 | 46.5 | 2.9 | 990 | 49.0 | 5.7 |
| Female | 2225 | 52.3 | 8.5 | 1193 | 53.5 | 6.2 | 1032 | 51.0 | 11.2 |
| Age | | | | | | | | | |
| 15-19 | 2365 | 55.6 | 4.7 | 1193 | 53.5 | 4.5 | 1172 | 58.0 | 4.9 |
| 20-24 | 1888 | 44.4 | 8.7 | 1038 | 46.5 | 4.9 | 850 | 42.0 | 13.4 |
| Marital Status | | | | | | | | | |
| Never Married | 3007 | 70.7 | 4.8 | 1401 | 62.8 | 3.9 | 1606 | 79.5 | 5.5 |
| Married | 1112 | 26.2 | 9.6 | 759 | 34.0 | 5.5 | 353 | 17.5 | 18.6 |
| Formerly married | 133 | 3.1 | 18.7 | 71 | 3.2 | 11.8 | 62 | 3.1 | 26.6 |
| Education | | | | | | | | | |
| No education | 181 | 4.2 | 7.6 | 147 | 6.6 | 4.9 | 34 | 1.7 | 19.1 |
| Primary | 1935 | 45.5 | 6.2 | 1412 | 63.3 | 4.2 | 523 | 25.9 | 11.6 |
| Secondary | 2016 | 47.4 | 6.6 | 663 | 29.7 | 5.7 | 1353 | 66.9 | 7.1 |
| Highest | 121 | 2.9 | 6.5 | 10 | 0.4 | - | 112 | 5.5 | 7.1 |
| Wealth | | | | | | | | | |
| low | 1291 | 30.4 | 3.7 | 733 | 32.8 | 4.0 | 412 | 20.4 | 8.8 |
| Medium | 1219 | 28.7 | 7.0 | 733 | 32.9 | 3.5 | 732 | 36.2 | 10.9 |
| High | 1743 | 41.0 | 8.1 | 766 | 34.3 | 6.5 | 878 | 43.5 | 6.2 |
| Employment | | | | | | | | | |
| Not employed | 2331 | 54.8 | 6.8 | 1005 | 45.1 | 5.9 | 1326 | 65.6 | 7.5 |
| Employed | 1921 | 45.2 | 6.1 | 1226 | 54.9 | 3.7 | 695 | 34.4 | 10.3 |
| Occupation | | | | | | | | | |
| Unemployed | 2332 | 54.8 | 6.8 | 1005 | 45.1 | 5.9 | 1326 | 65.6 | 7.5 |
| Professional | 598 | 14.1 | 11.9 | 169 | 7.6 | 10.0 | 429 | 21.2 | 12.7 |
| Agricultural | 978 | 23.0 | 3.0 | 904 | 40.5 | 3.0 | 73 | 3.6 | 2.7 |
| Manual | 216 | 5.1 | 6.7 | 73 | 3.3 | 1.1 | 143 | 7.1 | 9.6 |
| SEP Composite | | | | | | | | | |
| low | 1435 | 33.7 | 4.5 | 844 | 37.8 | 4.2 | 769 | 38.1 | 10.5 |
| Medium | 1471 | 34.6 | 8.0 | 1012 | 45.4 | 4.4 | 947 | 46.8 | 6.5 |
| High | 1347 | 31.7 | 6.9 | 375 | 16.8 | 6.4 | 306 | 15.1 | 9.5 |
| Religion | | | | | | | | | |
| Catholic | 904 | 21.3 | 5.12 | 447 | 20.0 | 4.52 | 457 | 22.6 | 5.71 |
| Protestant | 3270 | 76.9 | 6.90 | 1749 | 78.4 | 4.76 | 1521 | 75.3 | 9.36 |
| Muslim | 23 | 0.6 | 2.50 | 7 | 0.3 | - | 16 | 0.8 | 3.61 |
| Other | 46 | 1.1 | 2.58 | 26 | 1.1 | 4.65 | 21 | 1.0 | - |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Neighbourhood variables** | | | | | | | | | |
| Cluster Education | | | | | | | | | |
| low | 1191 | 28.0 | 4.7 | 686 | 30.8 | 2.5 | 646 | 32.0 | 11.7 |
| Medium | 1471 | 34.6 | 7.2 | 719 | 32.2 | 6.7 | 587 | 29.0 | 7.2 |
| High | 1591 | 37.4 | 7.2 | 826 | 37.0 | 4.8 | 788 | 39.0 | 6.8 |
| Cluster Wealth | | | | | | | | | |
| low | 1369 | 32.2 | 3.71 | 732 | 32.8 | 3.56 | 430 | 21.3 | 9.38 |
| Medium | 1137 | 26.7 | 7.20 | 799 | 35.8 | 4.05 | 675 | 33.4 | 10.72 |
| High | 1746 | 41.1 | 8.20 | 701 | 31.4 | 6.60 | 916 | 45.3 | 6.38 |
| Cluster Employment | | | | | | | | | |
| low | 1545 | 36.3 | 7.77 | 722 | 32.4 | 6.49 | 754 | 37.3 | 7.54 |
| Medium | 1382 | 32.5 | 6.82 | 815 | 36.5 | 5.02 | 709 | 35.1 | 7.72 |
| High | 1325 | 31.2 | 4.64 | 694 | 31.1 | 2.42 | 558 | 27.6 | 10.68 |
| Cluster SES | | | | | | | | | |
| low | 1353 | 31.8 | 3.87 | 732 | 32.8 | 3.40 | 480 | 23.7 | 10.59 |
| Medium | 1196 | 28.1 | 6.75 | 751 | 33.7 | 4.73 | 649 | 32.1 | 9.62 |
| High | 1704 | 40.1 | 8.38 | 749 | 33.6 | 5.91 | 892 | 44.2 | 6.49 |
| Total | 4253 | | 6.5 | 2231 | 100.0 | 4.7 | 2021 | 100.0 | 8.5 |

Table 2: Bivariate and Multivariate multilevel regression analysis of education as an underlying SEP risk factor for HIV infection among young people 15-24 years, by residence in Zambia

**First Approach (Education as the SEP indicator)**

| | Rural | | | | | | | | | Urban | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Bivariate | | Multivariate | | | | | | | Bivariate | | Multivariate | | | | | |
| **Predictor Variables** | | | | Model 1 | | Model 2 | | Model 3 | | | | | Model 1 | | Model 2 | | Model 3 | |
| | **Number** | **AOR** | **CI** | **OR** | **CI** | **OR** | **CI** | **OR** | **CI** | **Number** | **AOR** | **CI** | **OR** | **CI** | **OR** | **CI** | **OR** | **CI** |
| **Individual Variables** | | | | | | | | | | | | | | | | | | |
| SEX | | | | | | | | | | | | | | | | | | |
| Male | 1038 | 1 | Ref | | | | | 1 | Ref | 990 | 1 | Ref | | | | | 1 | Ref |
| Female | 1193 | 2.27 | (1.45 - 3.55)*** | | | | | 2.27 | (1.39 - 3.71)*** | 1032 | 2.68 | (1.89 - 3.81)*** | | | | | 2.36 | (1.63 - 3.43)*** |
| Age | 2231 | | | 1.08 | (1.01 - 1.16)** | 1.09 | (1.01 - 1.16)** | 1.06 | (0.97 - 1.16) | 2021 | | | 1.28 | (1.21 - 1.36)*** | 1.28 | (1.21 - 1.36)*** | 1.23 | (1.15 - 1.32)*** |
| Marital Status | | | | | | | | | | | | | | | | | | |
| Never Married | 1401 | 1 | Ref | | | | | 1 | Ref | 1606 | 1 | Ref | | | | | 1 | Ref |
| Married | 759 | 1.39 | (0.82 - 2.35) | | | | | 1.08 | (0.61 - 1.94) | 353 | 2.15 | (1.48 - 3.12)*** | | | | | 1.40 | (0.92 - 2.12) |
| Former married | 71 | 3.25 | (1.34 - 7.89)*** | | | | | 2.24 | (0.88 - 5.71)* | 62 | 3.20 | (1.73 - 5.91)*** | | | | | 2.07 | (1.07 - 3.97)** |
| Education in years | 2231 | 1.06 | (0.98 - 1.13) | 1.06 | (0.98 - 1.13) | 1.03 | (0.95 - 1.11) | 1.07 | (0.99 - 1.16) | 2021 | 0.93 | (0.88 - 0.98)*** | 0.93 | (0.88 - 0.98)*** | 0.93 | (0.88 - 0.98)*** | 0.96 | (0.90 - 1.01) |
| **Neighbourhood Variables** | | | | | | | | | | | | | | | | | | |
| Cluster Education | | | | | | | | | | | | | | | | | | |
| low | 686 | 1 | Ref | | | 1 | Ref | 1 | Ref | 646 | 1 | Ref | | | 1 | Ref | 1 | Ref |
| Medium | 719 | 2.88 | (1.54 - 5.38)*** | | | 2.74 | (1.45 - 5.20)*** | 2.57 | (1.35 - 4.88)*** | 587 | 0.79 | (0.52 - 1.21) | | | 0.89 | (0.57 - 1.37) | 0.87 | (0.56 - 1.36) |
| High | 826 | 2.04 | (1.08 - 3.84)** | | | 1.86 | (0.94 - 3.66)* | 1.72 | (0.87 - 3.40) | 788 | 0.72 | (0.47 - 1.10) | | | 0.90 | (0.57 - 1.44) | 0.93 | (0.57 - 1.50) |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Fixed effect model test** | | | | | | |
| Wald | | 16.14 | 33.33 | | 74.06 | 105.79 |
| **Variance of random effects** | | | | | | |
| Variance | | 0.46 | 0.47 | | 0.20 | 0.20 |
| **Model Statistics** | | | | | | |
| $R^2$ | | 0.19 | 0.18 | | 0.01 | -0.01 |
| Bivariate & Model 2 | | | | | | |
| Log likelihood test | 10.16 | | | 0.32 | | |
| Prob > chi2 | 0.01 | | | 0.85 | | |

*P<0.10, **P<0.05, ***P<0.01
OR: Odds Ratio
AOR: Age adjusted Odds Ratio
CI Confidence Intervals

Table 3: Bivariate and Multivariate multilevel regression analysis of education, wealth and employment as underlying SEP risk factors for HIV infection among young people 15-24 years, by residence in Zambia

**Second Approach (All SEP Indicators)**

| | Rural | | | | | | | | | Urban | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Bivariate | | | Multivariate | | | | | | Bivariate | | | Multivariate | | | | | |
| **Predictor Variables** | | | | Model 1 | | Model 2 | | Model 3 | | | | | Model 1 | | Model 2 | | Model 3 | |
| | Number | AOR | CI | OR | CI | OR | CI | OR | CI | Number | AOR | CI | OR | CI | OR | CI | OR | CI |
| **Individual Variables** | | | | | | | | | | | | | | | | | | |
| Sex | | | | | | | | | | | | | | | | | | |
| Male | 1038 | 1 | Ref | | | | | 1 | Ref | 990 | 1 | Ref | | | | | 1 | Ref |
| Female | 1193 | 2.27 | (1.45 - 3.55)*** | | | | | 2.16 | (1.29 - 3.62)*** | 1032 | 2.68 | (1.89 - 3.81)*** | | | | | 2.30 | (1.56 - 3.38)*** |
| Age | 2231 | | | 1.12 | (1.04 - 1.21)*** | 1.12 | (1.04 - 1.20)*** | 1.08 | (0.98 - 1.18) | 2021 | | | 1.32 | (1.24 - 1.40)*** | 1.32 | (1.24 - 1.40)*** | 1.25 | (1.17 - 1.34)*** |
| Marital Status | | | | | | | | | | | | | | | | | | |
| Never Married | 1401 | 1 | Ref | | | | | 1 | Ref | 1606 | 1 | Ref | | | | | 1 | Ref |
| Married | 759 | 1.39 | (0.82 - 2.35) | | | | | 1.15 | (0.64 - 2.09) | 353 | 2.15 | (1.48 - 3.12)*** | | | | | 1.33 | (0.87 - 2.02) |
| Former married | 71 | 3.25 | (1.34 - 7.89)*** | | | | | 2.48 | (0.96 - 6.39) | 62 | 3.20 | (1.73 - 5.91)*** | | | | | 2.03 | (1.05 - 3.92)** |
| Education in years | 2231 | 1.05 | (0.98 - 1.13) | 1.02 | (0.95 - 1.10) | 1.00 | (0.92 - 1.08) | 1.04 | (0.96 - 1.13) | 2021 | 0.93 | (0.88 - 0.98)*** | 0.92 | (0.88 - 0.97)*** | 0.93 | (0.88 - 0.98)*** | 0.95 | (0.90 - 1.01)* |
| Wealth Categories | | | | | | | | | | | | | | | | | | |
| low | 733 | 1 | Ref | 1 | Ref | | | 1 | Ref | 412 | 1 | Ref | | | | | 1 | Ref |
| Medium | 733 | 0.82 | (0.46 - 1.45) | 0.78 | (0.44 - 1.38) | 0.68 | (0.38 - 1.22) | 0.62 | (0.34 - 1.13) | 732 | 1.22 | (0.83 - 1.79) | 1.35 | (0.90 - 2.01) | 1.15 | (0.72 - 1.85) | 1.14 | (0.70 - 1.84) |
| High | 766 | 1.67 | (1.00 - 2.78)** | 1.51 | (0.89 - 2.57) | 1.21 | (0.67 - 2.21) | 1.21 | (0.66 - 2.22) | 878 | 0.78 | (0.51 - 1.19) | 0.93 | (0.59 - 1.46) | 0.85 | (0.47 - 1.53) | 0.83 | (0.46 - 1.53) |
| Employment | | | | | | | | | | | | | | | | | | |
| Not employed | 1005 | 1 | Ref | 1 | Ref | | | 1 | Ref | 1326 | 1 | Ref | | | | | 1 | Ref |
| Employed | 1226 | 0.53 | (0.34 - 0.81)*** | 0.54 | (0.35 - 0.84)*** | 0.63 | (0.41 - 0.99)** | 0.78 | (0.49 - 1.24) | 695 | 0.65 | (0.46 - 0.91)*** | 0.63 | (0.45 - 0.89)*** | 0.63 | (0.44 - 0.88)*** | 0.81 | (0.56 - 1.16) |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Neighbourhood Variables** | | | | | | | | | | | | | | | | |
| Cluster Education | | | | | | | | | | | | | | | | |
| low | 686 | 1 | Ref | | | | 1 | Ref | 646 | 1 | Ref | | | | 1 | Ref |
| Medium | 719 | 2.88 | (1.54 - 5.38)$^{***}$ | | 2.79 | (1.46 - 5.31)$^{***}$ | 2.71 | (1.42 - 5.17)$^{***}$ | 587 | 0.79 | (0.52 - 1.21) | | 0.76 | (0.47 - 1.21) | 0.73 | (0.45 - 1.17) |
| High | 826 | 2.04 | (1.08 - 3.84)$^{**}$ | | 1.58 | (0.77 - 3.27) | 1.46 | (0.70 - 3.03) | 788 | 0.72 | (0.47 - 1.10) | | 0.90 | (0.48 - 1.70) | 0.87 | (0.46 - 1.67) |
| Cluster Wealth | | | | | | | | | | | | | | | | |
| low | 732 | 1 | Ref | | | | 1 | Ref | 430 | 1 | Ref | | | | 1 | Ref |
| Medium | 799 | 1.18 | (0.64 - 2.19) | | 0.96 | (0.51 - 1.84) | 0.94 | (0.49 - 1.79) | 675 | 1.35 | (0.89 - 2.05) | | 1.54 | (0.90 - 2.65) | 1.66 | (0.95 - 2.89) |
| High | 701 | 1.98 | (1.11 - 3.52)$^{**}$ | | 1.19 | (0.59 - 2.38) | 1.17 | (0.58 - 2.36) | 916 | 0.82 | (0.53 - 1.29) | | 1.17 | (0.53 - 2.56) | 1.31 | (0.59 - 2.93) |
| Cluster Employment | | | | | | | | | | | | | | | | |
| low | 722 | 1 | Ref | | | | 1 | Ref | 754 | 1 | Ref | | | | 1 | Ref |
| Medium | 815 | 0.74 | (0.44 - 1.23) | | 0.82 | (0.49 - 1.37) | 0.77 | (0.46 - 1.29) | 709 | 0.96 | (0.62 - 1.48) | | 1.01 | (0.65 - 1.56) | 0.98 | (0.63 -1.52) |
| High | 694 | 0.34 | (0.18 - 0.63)$^{***}$ | | 0.41 | (0.21 - 0.78)$^{***}$ | 0.38 | (0.20 - 0.73)$^{***}$ | 558 | 0.99 | (0.64 - 1.53) | | 1.05 | (0.66 - 1.67) | 1.03 | (0.65 - 1.65) |
| **Fixed effect model test** | | | | | | | | | | | | | | | | |
| Wald | | | | 20.84 | 36.89 | | 51.16 | | | | | 84.64 | 87.20 | | 113.87 | |
| **Variance of random effects** | | | | | | | | | | | | | | | | |
| Variance | | | | 0.52 | 0.31 | | 0.31 | | | | | 0.17 | 0.16 | | 0.17 | |
| **Model Statistics** | | | | | | | | | | | | | | | | |
| $R^2$ | | | | 0.08 | 0.46 | | 0.46 | | | | | 0.15 | 0.19 | | 0.17 | |
| Model 1 & Model 3 | | | | | | | | | | | | | | | | |
| Log likelihood test | | | | 19.67 | | | | | | 3.77 | | | | | | |
| Prob > chi2 | | | | 0.00 | | | | | | 0.71 | | | | | | |

$^{*}$P<0.10, $^{**}$P<0.05, $^{***}$P<0.01, OR: Odds Ratio, AOR: Age adjusted Odds Ratio, CI Confidence Intervals

Table 4: Bivariate and Multivariate multilevel regression analysis of a SEP composite indicator as an underlying SEP risk factor for HIV infection among young people 15-24 years, by residence in Zambia

**Third Approach (Composite SES)**

| | Rural | | | | | | | | | Urban | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Bivariate | | Multivariate | | | | | | | Bivariate | | Multivariate | | | | | |
| **Predictor Variables** | | | | Model 1 | | Model 2 | | Model 3 | | | | | Model 1 | | Model 2 | | Model 3 | |
| | | AOR | CI | OR | CI | OR | CI | OR | CI | | AOR | CI | OR | CI | OR | CI | OR | CI |
| **Individual Variables** | | | | | | | | | | | | | | | | | | |
| Sex | | | | | | | | | | | | | | | | | | |
| Male | 1038 | 1 | Ref | | | | | 1 | Ref | 990 | 1 | Ref | | | | | 1 | Ref |
| Female | 1193 | 2.27 | (1.45 - 3.55)*** | | | | | 2.26 | (1.38 - 3.69)*** | 1032 | 2.68 | (1.89 - 3.81)*** | | | | | 2.26 | (1.56 - 3.29)*** |
| Age | 2231 | | | 1.08 | (1.01 - 1.16)** | 1.08 | (1.01 - 1.16)** | 1.06 | (0.97 - 1.16) | 2021 | | | 1.29 | (1.21 - 1.37)*** | 1.29 | (1.22 - 1.37)*** | 1.23 | (1.15 - 1.32)*** |
| Marital Status | | | | | | | | | | | | | | | | | | |
| Never Married | 1401 | 1 | Ref | | | | | 1 | Ref | 1606 | 1 | Ref | | | | | 1 | Ref |
| Married | 759 | 1.39 | (0.82 - 2.35) | | | | | 1.00 | (0.56 - 1.78) | 353 | 2.15 | (1.48 - 3.12)*** | | | | | 1.47 | (0.98 - 2.22)* |
| Former married | 71 | 3.25 | (1.34 - 7.89)*** | | | | | 2.09 | (0.82 - 5.31) | 62 | 3.20 | (1.73 - 5.91)*** | | | | | 2.23 | (1.18 - 4.21)*** |
| SEP Composite | | | | | | | | | | | | | | | | | | |
| low | 844 | 1 | Ref | 1 | Ref | 1 | Ref | 1 | Ref | 769 | 1 | Ref | | | | | 1 | Ref |
| Medium | 1012 | 1.01 | (0.63 - 1.62) | 1.01 | (0.63 - 1.62) | 0.86 | (0.53 - 1.42) | 0.95 | (0.58 - 1.57) | 947 | 0.68 | (0.48 - 0.96)** | 0.68 | (0.48 - 0.96)** | 0.63 | (0.44 - 0.92)** | 0.75 | (0.51 - 1.11) |
| High | 375 | 1.41 | (0.80 - 2.51) | 1.41 | (0.80 - 2.51) | 1.08 | (0.57 - 2.05) | 1.33 | (0.69 - 2.57) | 306 | 0.50 | (0.30 - 0.82)*** | 0.50 | (0.30 - 0.82)*** | 0.46 | (0.27 - 0.79)*** | 0.59 | (0.34 - 1.04)* |
| **Neighbourhood Variables** | | | | | | | | | | | | | | | | | | |
| Cluster SEP Composite | | | | | | | | | | | | | | | | | | |
| low | 732 | 1 | Ref | | | 1 | Ref | 1 | Ref | 480 | 1 | Ref | | | | | 1 | Ref |
| Medium | 751 | 1.55 | (0.84 - 2.87) | | | 1.58 | (0.84 - 2.97) | 1.52 | (0.81 - 2.86) | 649 | 1.18 | (0.78 - 1.79) | | | 1.42 | (0.92 - 2.18) | 1.48 | (0.96 - 2.29) |
| High | 749 | 1.90 | (1.05 - 3.45)** | | | 1.89 | (0.97 - 3.68)* | 1.75 | (0.90 - 3.43) | 892 | 0.83 | (0.54 - 1.29) | | | 1.23 | (0.75 - 2.03) | 1.25 | (0.75 - 2.09) |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Fixed effect model test** | | | | | | |
| Wald | | 10.05 | 26.87 | | 76.72 | 109.18 |
| **Variance of random effects** | | | | | | |
| Variance | | 0.53 | 0.54 | | 0.17 | 0.16 |
| **Model Statistics** | | | | | | |
| $R^2$ | | 0.07 | 0.06 | | 0.17 | 0.19 |
| Bivariate & Model 3 | | | | | | |
| Log likelihood test | 3.74 | | | 2.50 | | |
| Prob > chi2 | 0.15 | | | 0.29 | | |

*P<0.10, **P<0.05, ***P<0.01
OR: Odds Ratio
AOR: Age adjusted Odds Ratio
CI Confidence Intervals