The social epidemiology of HIV infection: a study among unmarried young people in rural South Africa in 2001

James R. Hargreaves

London School of Hygiene and Tropical Medicine

November 2005

Thesis submitted to the University of London for the degree of Doctor of Philosophy

Abstract

In sub-Saharan Africa, risk of HIV infection was associated with increased education, mobility and income in studies completed before 1996. Some later studies found no association or a negative socioeconomic gradient in prevalent HIV infection. The social epidemiology of HIV infection among South African young people remains poorly understood.

This study examined the association between socioeconomic factors and HIV risk among unmarried young people aged 14-25 years in rural South Africa in 2001. Some 916 males (73.6% of those eligible) and 1003 females (80.8%) were included in the study sample.

Households were ranked into those that were “very poor”, “poor, but a bit better off” or “doing OK”. Most young people attended school, though this was associated with increased wealth among the females. Those not attending school were mostly unemployed. Temporary migration was increasingly common with greater age but was not associated with household wealth. Low HIV awareness and risky sexual behaviour characteristics were widespread. HIV prevalence was 5.6% among males and 12.2% among females.

Household wealth was not associated with HIV risk except that young women from wealthier households more often reported condom use. Those attending school of both sexes reported lower numbers of sexual partners than those not attending. Young women attending school reported partnerships with lower frequency of intercourse, men closer in age to themselves and higher levels of condom use. HIV prevalence was lowest among young men attending school. Temporary migrants of both sexes reported greater levels of HIV awareness than those staying at home, while female migrants also reported more condom use.

In this study, young people were at high risk of HIV infection, particularly those not attending school. Young people remaining in rural areas were also at high risk of new infection. Programmes that reach these young people and increase educational opportunities might help reduce HIV infection rates.

Contents

List of tables and figures

Acknowledgements

The list of people who have made massive contributions to the process of my writing this thesis is so long as to be quite embarrassing. I owe an enormous debt to so many people without whom I simply wouldn't have got to the point where this document exists. I hope this work, the mistakes in which are all mine, does them all justice. Forgive me if I ramble a little here ...

Linda Morison's wisdom, kindliness and support have been a constant feature of the process of preparing and writing this thesis and in my life more generally over the past six (!) years. My admiration for her, and my level of debt to her, are immense. I can't really think of anything else to say.

Without the energy and vision of Paul Pronyk the IMAGE study would not exist. I continue to be inspired by Paul and hope that the research we continue to do together might, in some modest way, be contributing to the fight against HIV/AIDS. I am glad of an opportunity to express my thanks. Similarly, Julia Kim, Benjamin Makhubele and latterly Godfrey Phetla and Rico Euripidou, who made up the rest of the research team that planned and implemented the IMAGE study have been inspirational over many years. John Gear and Mmatshilo Motsei both hold special places in that long story, while John Porter and Charlotte Watts made things happen in London and have contributed much to my learning along the way. Dave Leon and Sarah Salway have contributed valuably to my thinking as members of my advisory panel. The IMAGE Study would also not have happened without the generous support of the following donors - AngloAmerican Chairman's Educational Trust, AngloPlatinum, Department for International Development (UK), The Ford Foundation, The Henry J. Kaiser Family Foundation, HIVOS, South African Department of Health and Welfare, and the Swedish International Development Agency.

My fortune in working with some of the most wonderful people astounds me. Joseph Mhlaba, Lulu (Xirandzhiwa) Ndhlovu and Venice Mbowane all made the journey out to Sekhukhuneland with me and set up their lives there to assist in collecting the data reported in this thesis. Edwin Maroga joined later; his work as data manager for the study has been invaluable and made my life immeasurably easier. The field team who collected the survey data presented here and put up with me through four long, dusty months in 2001 (and in many cases much longer) were Madihlare Kgwete, Kedibone Mabuza, Julia Mathipa (nee Sekgobela), Alinah Magopane, Malebo Nkuna, Leah Makhubedu, Juliet Makofane, Leah Makofane, Anna Mampho, Maria Malepe, Nomsa Mgiba, Junior Nkadimeng and Fridah Sello. A million thanks are due to them all for their hard work, support and friendship and most of all for bestowing me a new name.

Staff of the Small Enterprise Foundation who conducted participatory wealth ranking over such a wide area in such a short time were Johannes Matsika, Germina Ledwaba, Gabriel Rapatsa, Oxygen Rivombo, Evah Motshwane, Alfrida Ramoroka and Noria Manganyi. Support from John de Wit, Ben Nkuna, Brian Kuwik and Kalipe Mashaba was essential for the project to be a success – and it has never wavered.

In Johannesburg, largely hidden from the day-to-day trials of our work in Sekhukhuneland, Gwynn Stevens, Anusha Makhuraj, Wendy Stevens and later Grant Napier, analysed the biological specimens collected in the field and compiled the laboratory data held in this thesis – getting themselves inadvertently embroiled in our story along the way.

Two students from USA helped greatly in aspects of the work I report here; Katherine Rowe was a joy to have as support during the survey fieldwork and checked many hundreds of questionnaires. Chris Martin spent many long hours coding statements from the PWR process into a database. Many friends at the Wits Rural Facility where I lived from 2000-2004 played a part in getting me to the end of this road. Neil helped type many tables, many times. Chris Bonell and my dad both read this thesis and gave incredibly useful comments - cheers. I'm sure there are many many others who I've neglected to mention.

Finally, during the long, tortuous days while I have been pulling this thesis together, Helen has put up with my moaning and kept me sane with endless love and laughs. I can't thank her enough. X.

dedicated to the people who have generously given their valuable time to participate in the IMAGE study

List of Acronyms

AIDS	Acquired Immune Deficiency Syndrome
ANC	Antenatal clinic
CI	Confidence Interval
DOH	Department of Health
DHS	Demographic and Health Surveys
ELISA	Enzyme-linked immunosorbent assay
GBV	Gender Based Violence
GPA	Global Programme on AIDS
HIV	Human Immunodeficiency Virus
HSRC	Human Sciences Research Council
IMAGE	Intervention with Microfinance for AIDS and Gender Equity
IPV	Intimate Partner Violence
IQ	Interquartile
KZN	KwaZulu Natal
LSHTM	London School of Hygiene and Tropical Medicine
MFI	Microfinance Initiative
NHLS	National Health Laboratory Service
OMT	Oral Mucosal Transudate
OR (aOR)	Odds Ratio (Adjusted Odds Ratio)
PCA	Principal Components Analysis
PEP	Post Exposure Prophylaxis
PhD	Doctorate of Philosophy
PRA	Participatory Rural Appraisal
PWA	Person living with HIV/AIDS
PWR	Participatory Wealth Ranking
RADAR	Rural AIDS and Development Action Research Programme
SA	South Africa
SEF	Small Enterprise Foundation
SES	Socioeconomic status
SFL	Sisters for Life
STI	Sexually Transmitted Infection
TCP	Tšhomišano Credit Programme
UNAIDS	United Nations Joint Programme on AIDS
VCT	Voluntary Counselling and Testing for HIV

Glossary

Term	Definition
Population under study	
Household	A group of people who are residents at the same dwelling, who may currently be either staying at home or away from the home at the time of the survey, but who eat from the same pot of food when staying there.
Adolescent	An individual in the life period between puberty and adulthood.
Young adult	An individual in the early years of adulthood
Young people	For the purposes of this thesis, "young people" denotes adolescents and young adults aged 14-25 years
Marriage	Used to refer to those in partnerships characterised by a formal marriage ceremony, either Christian or traditional. *Lobola* (bride-price) may have been paid, partly paid or may not have been paid at all. Unmarried individuals, who are the focus of this thesis, were those living singly who were not in such partnerships.
Socioeconomic factors	
Socioeconomic position	Position in distributions of, for example, wealth, income or education.
Relative household wealth	The marker of socioeconomic position used in this study. Denotes the relative level of welfare experienced by a household unit that is primarily the function of its socioeconomic activities. This distinguishes the concept from other dimensions of welfare such as happiness, spirituality or disability.
Socioeconomic role	Socioeconomic activities undertaken by study participants during the past year including attendance at school/college and temporary migrancy.
School or college attendance	Denotes full-time attendance at a school or college, not including enrolment in part time courses, non-attendance courses, work related courses, evening classes or other educational courses not requiring attendance.
Temporary migrant	Refers to individuals considered members of a rural household, but who for significant periods of time are not sleeping within the household dwelling. They may be attending school or may be looking for or taking up work for pay at a location distant from the rural household. The terms circular or labour migrancy have also been used to describe this phenomenon.
Sexual behaviour	
Sexual intercourse	Penetrative vaginal intercourse within heterosexual partnerships. Although data collection tools used in this study did not actively make distinctions on the basis of sex of partner or type of intercourse, discussions with fieldworkers suggested that reporting of homosexual partners or anal intercourse were likely to be rare.
Epidemiological terms	
Exposure	The putative causal agent under investigation in an epidemiological analysis.
Outcome	The variable hypothesised to be causally related to the exposure variable in an epidemiological analysis.
Confounder	Confounding variables mask the true association between an exposure and outcome, and must be appropriately taken into consideration in statistical analyses. For a variable to act as a confounder, it must be independently associated with the outcome variable in question, associated with the exposure, must not be a pathway variable and statistical adjustment for the factor will alter the measure of association between exposure and outcome.
Pathway variable	Variables hypothesised to lie on the causal chain between a postulated exposure and outcome. Also termed *mediator* variables.
Interaction	The phenomenon of the magnitude and / or direction of association between an exposure and an outcome being dependent on some other factor. Also known as effect modification.

Chapter 1. Introduction

1.01 Adolescence and young adulthood: a time of development

Children are almost entirely dependent upon those around them, in particular their parents or surrogate guardians, and in most societies this dependence continues throughout childhood. The onset of adolescence sees the starting point from which the dependence of individuals on others begins to recede as the journey towards full adulthood begins.

Typically, adolescence and young adulthood is a period of healthiness un-matched throughout the life course (1, 2). Mortality and morbidity rates in this age group are typically lower than those seen both in childhood and in later life. Nevertheless, this period of life is associated with particular issues that have the potential to influence health, both during adolescence and young adulthood, but also later in life. Adolescent experimentation with alcohol, smoking or drug intake may precipitate longer term habits with important health consequences (3). Adolescents may also be at particularly high risk of violent and accidental death, from causes such as suicide, criminal violence and car accidents. The onset of sexual behaviour also brings with it important health issues. Early childbearing among young women can influence their health directly, but might also have longer term social and economic consequences. In many countries of the world today, the risk of acquisition of STIs and, in particular, HIV/AIDS also pose a particular health risk generally first encountered during adolescence.

This thesis presents data from an epidemiological study conducted in rural South Africa. It will describe the association between socioeconomic factors and the sexual behaviour and risk of HIV infection of unmarried young people in rural South Africa. The setting for the study was the Sekhukhuneland region of South Africa's north-east, on the border of the Limpopo and Mpumalanga provinces in 2001.

1.02 Structure of the thesis

Chapter 2 will provide a background to the study in three sections. Firstly, it will describe the socio-economic environment and prevailing patterns of sexual behaviour and HIV infection in South Africa. Secondly, it will discuss the theoretical basis for the current study. The chapter concludes with the development of a conceptual framework to guide the research presented in this thesis.

Chapter 3 will present the results of a review of the empirical literature on the social epidemiology of HIV risk characteristics from sub-Saharan Africa. It concludes by identifying trends and gaps in the literature and developing hypotheses to be tested in the current research.

The remainder of the thesis will describe the aims, methods, results and implications of the current study. Chapter 4 will outline the objectives of the study, provide a brief description of the history and current context of the study region and describe the methods of data collection and analysis.. Chapter 5 will describe the characteristics of households and individuals eligible for and included in the study, and discuss factors related to non-inclusion in the final sample. Chapter 6 will present data on wealth, school attendance and temporary migrancy among the study sample. The chapter will describe the generation of three indices of household wealth and investigate the relationship between these measures, before one is selected for use in the remainder of the thesis. The chapter will go on to examine the individual socioeconomic roles played by young people and their association with household wealth. Chapter 7 will describe sexual behaviour characteristics reported by young people and the pattern of HIV infection among this group. Chapter 8 will describe the main associations pertaining to the study hypotheses. Firstly, it will present measures of the association between household wealth and risk of HIV infection. This will be followed by measures of association between the individual socioeconomic roles played by young people and characteristics of their sexual behaviour and risk of HIV infection. Finally, data will be presented on interactions between the socioeconomic exposures and their association with HIV risk outcomes, and on the potential role of individual socioeconomic roles in mediating any association between household wealth and HIV risk. Each results chapter will conclude with a brief summary of the main findings presented. Following this, chapter 9 will present analyses that seek to investigate the sensitivity of the main findings of the study to a variety of assumptions about the validity of the collected data and the potential for selection bias to exist. Finally, Chapter 10 will identify expected and unexpected findings of the study, discuss the strengths and limitations of the research and assess the findings of the study in light of the preceding empirical and theoretical literature, considering potential policy and programmatic implications of the study.

Chapter 2. Background

2.01 The context of adolescence and young adulthood in South Africa

The adolescent and young adulthood period is characterised by changes at both the individual and contextual level, which are driven by biological, cognitive and social processes (4). Early adolescence is characterised by puberty, a period of rapid physical development driven by major hormonal changes in the body. Physical changes include growth spurt, the onset of growth of pubic and other body hair, voice changes among males, menarche among females and the development of the sex organs and onset of sexual desire (5).

Early physical changes of puberty are accompanied and followed into adolescence by cognitive and psychosocial changes. Adolescence is characterised by cognitive development, and increased levels of autonomy and self identity (4). It is a period of experimentation with new ways of acting and new experiences. Adolescents increasingly see themselves in terms of their individuality and this catalyses change in the nature and function of their relationships with parents, family members and peer groups. Individual changes are accompanied by significant changes in the expectations of the roles to be played by young people. As young people begin to enter adulthood, laws and customs allow them to do things they were previously prohibited from doing, such as work for money, have sex, smoke tobacco or drink alcohol. The physical and emotional development of adolescence requires them not just to be dependent on their families or households, but to contribute to them. In this way, as adolescents approach full adulthood, the roles they are expected to play by those around them change. Adolescents might be increasingly expected to contribute to the household in the form of chores, work or money. Eventually, most young people assume the full roles of adulthood, such as becoming independent, leaving home, marrying, and having children themselves.

Young people in South Africa today are faced with many opportunities, but also with significant challenges. Among the most serious of these are the experience of household poverty, the transition from a developing educational system into a labour market characterised by high levels of unemployment and migration, and the negotiation and formation of sexual relationships in an era of HIV/AIDS.

2.01.1 Socio-economic context

This study focuses on individuals aged 14-25 years. Most of those at the upper end of this age spectrum were born in the year 1976, a momentous year for South Africa and her young people. On June 16 of that year, thousands of students took to the streets in Soweto to protest at laws that would affect the language and quality of their education. The day turned to bloodshed, and over the next three days over 170 people were killed. The 'Soweto uprising' was a turning point in the South African struggle and was followed by a period of social turmoil, intense youth activism and civil action that sowed the seeds of the end of apartheid (6, 7).

The 1980s and 1990s were marked by violence and unrest in South Africa, and were characterised by school boycotts and increasing alienation between youth "comrades" and their peers. By 1987, when the youngest of those included in this study were being born, and the eldest were entering puberty and the age of secondary education, school boycotts were rife, violence endemic and the South African authorities declared a state of emergency in the country. There was, at this time, little optimism and scant sign of the radical changes that would follow. Yet just three years later, Nelson Mandela was released from prison after 27 years on Robben Island. Widespread unrest continued but by 1994 Mandela was president and South Africa's transition to democracy was acknowledged the world over.

Miraculous as the transition was, however, the seven years between 1994 and 2001, the year in which the research project on which this thesis is based was conducted, were not easy. While South Africa internationally became a symbol of peace, little changed for much of the country's population. The post-apartheid years have been full of hope, but have also seen rises in unemployment, increases in inequality and crime, and the entrenchment of the most serious epidemic of HIV/AIDS in the world today. In 2001, South Africa is a democratic middle income country characterized by high levels of income inequality (8). The apartheid regime entrenched these inequalities along racial lines, with the African majority being most disadvantaged. This situation persists today.

Consequently, household poverty remains a profound reality for the majority of black South African youth. Since 1994, the ruling ANC party has pursued an economic agenda focused on international trade and black economic empowerment. These policies have supported South Africa's international reputation and have fostered the growth of a minority black elite, but have done little to impact the lives of the majority of poor households. Since the end of apartheid there has been a rise in inequality within racial groups (8). Yet the highest levels of poverty, in terms of both income and basic needs, remain to be found among the African population, with rural areas worst affected. Figures from studies conducted in the mid 1990s suggested that 43% of the national population, and 57% of the population of Limpopo province, where this study is set, were food insecure (9), and 61.7% of households and 74.9% of individuals in Limpopo province were in income poverty (10).

The roots of entrenched rural poverty in South Africa are to be found in the colonial and apartheid history of the country (11). Today, a number of factors sustain high levels of poverty in rural areas of South Africa. The former bantustans of South Africa, where much of the rural African population live, are overcrowded, unable to support agriculture to underpin household livelihoods and support little or no industry, making households heavily reliant on cash earned in towns (10, 12, 13). A study in 1995 suggested that the major sources of income for rural African households were remittance from migrant labour, local labour and state grants (11). However, in South Africa there is high, and rising, unemployment, even in urban areas, with 30% of adults in 2002 unemployed by a definition which includes only those actively seeking work (41% were unemployed by a broader definition not requiring work seeking) (14). Increases in the cost of labour following the end of apartheid precipitated massive downsizing of major employment industries, notably mining and manufacturing. This has resulted in high levels of unemployment and reductions in household income for many. Alongside this, there are significant barriers to entry into the informal sector, including crime, lack of access to credit, poor infrastructure and the psychosocial legacy of apartheid (15, 16). It has been suggested that much of the rural population is caught in a 'poverty trap' with little means of escape (17).

The government has prioritised a number of aspects of poverty alleviation and made significant progress in these regards. Most notably this has been in the provision of basic services (particularly water, electricity and telecommunications) and housing. Large government programmes in these arenas have made significant improvements to the lives of many, and the government continues to target poverty alleviation as a central component of its agenda. Nevertheless, there remains much to do to improve the standard of living of the majority of rural South African households.

In this environment, emerging from childhoods immersed in the household, the progression to adulthood for many young South Africans is characterised by the move from school to labour market. Young people may continue in education until the completion of secondary school ("matriculation") or enter further education, or they may leave school and enter the labour market. They might also migrate for a number of reasons, including to look for or to take up work, to lessen the burden on the rural homestead, to seek better opportunities for education or other reasons, or they may remain in the rural home. The roles that individuals play reflect a balance of immediate and long term needs, and individual, household and external forces. The following paragraphs will discuss the context of these socioeconomic roles in rural South Africa.

The apartheid regime's policy of Bantu Education was one of the cornerstones of its ideology and one of the most hated and damaging of all of its programmes. While progress has been made since 1994, significant challenges to the education of young people remain. The South African education system is characterized by almost universal enrolment at primary school ages, high levels of enrolment at ages beyond the standard leaving age of 18, high levels of literacy, but low levels of completion of secondary school ("matriculation") (18, 19). The system often sees students repeat years or drop out of school for short periods (18, 20), and it has been suggested that this, as opposed to non-enrolment per se, is the major reason for the limited secondary school completion rate (18). Reasons for non-advancement include lack of school fees, illness, a need to work, and, for young women, pregnancy (20, 21). The apartheid system entrenched a low quality of schooling in rural areas, and despite significant attempts to improve school quality since 1994, many rural schools remain dirty, noisy, overcrowded, with low quality of teaching, and high levels of corporal punishment and sexual abuse of female students (20).

Despite these issues, schooling is seen as highly important by many South African youth and their families (21). Adults with higher levels of education, and in particular those who complete secondary school or beyond, are more likely to get a job and to earn higher wages in the labour market (18, 22). Educated labour has been suggested to be the most important livelihood asset for rural households (10). Yet many households also view the work potential of students in contributing to the household economy as important, with 11-12% of both male and female students aged 19-24 working while they were at school in one study in KwaZulu Natal (20). This work, alongside a high burden of household chores, particularly for young women, is likely to provide constraints to successful schooling. Alternatively, continued investment in schooling may be judged too great a cost by households and individuals, and young people may be encouraged to enter the labour market.

The apartheid system was precisely set up to provide almost no employment opportunities in the bantustans, ensuring an easily regulated and cheap source of labour migrants for industry. Local employment opportunities in many former homelands remain low or virtually non-existent, and the labour market remains centralised to urban areas. Consequently, unemployment in South Africa is highest among rural populations, among women and among youth (22). The rate of adult unemployment in 2002 in South Africa was 30% and 41%, using the narrow and broad definitions. However, unemployment among youth was even higher than this, and the highest rates seen in South Africa are in Limpopo Province (23). Youth unemployment is seen as one of the biggest problems affecting South Africa today, and has been linked to wider problems in society such as crime, drug and alcohol use, and male suicide (24). Literature on unemployment in South Africa has sought to explain the high persistent levels seen. It has been shown that compared with other countries, relatively few individuals, particularly young people, enter self employment, even though this would be likely to provide higher income (16). Nevertheless, unemployment in South Africa is largely involuntary, with barriers to entry into self employment (such as crime and lack of credit) providing the reasons that so few seek this. Other authors have pointed to aspects of the socioeconomic environment that create low job prospects, and have highlighted the poor education system, poor infrastructure, a dysfunctional health service and high levels of crime as being of central importance (25). It remains unclear what the solution to the problem might be. Most authors agree that investment in training and education, and removing barriers to informal market entry are likely to be important (21, 26). As discussed above, young people with higher levels of educational attainment, and those with any prior job experience are those least likely to remain unemployed.

Migration to find work remains an entrenched component of rural life in South Africa, with remittances from household members sleeping away from the rural home being the most important source of income for many rural households (10). While some households and individuals permanently relocate, a majority of individuals retain links with a rural home, with many seeing this as their "true" home (27, 28). During apartheid, with its restrictions on movement for the black population, a prevailing pattern for rural homes was for male adults of the household, particularly the head, to work away from the home throughout much of the year, with mines and big industry being important sources of employment. Generally these men regularly sent money back to support the rural home and returned home only a few times during the year. Since apartheid much of this system has remained, although the profile of migrants is changing, with increased levels of migration among females and young people, higher levels of permanent migration, and higher levels of migration to local towns (29).

Young people might be temporarily migrant from the rural home for a number of reasons. Young people may reside away from the rural home in an area with better access to schooling (30). Alternatively, young people who have left school may leave the rural area to find or take up work elsewhere. The destination may be both known and friendly, or neither of these. Temporary migrants from rural areas are more likely to be in employment than those who remain at home (29, 31). Levels of migration among young people are generally higher than for children or for adults approaching or beyond pension age, but lower than that for older adults. For example, in the Agincourt sub-district of Limpopo Province, approximately one third of 15-34 year old males were not sleeping in the home for more than six months of the previous year, compared to nearly two thirds of 35-65 year olds. Among females the corresponding figures were approximately 15% and 20% (29).

In summary, while the period since apartheid has seen large scale investment in infrastructure, with improvements in access to water, housing, telephones, electricity and sanitation, all of which have contributed to an improved standard of living, this period has also seen a rise in unemployment and falls in household income, and much of the rural population remains in poverty today. Most young people of both sexes attend primary and some secondary school, but relatively few complete secondary school due to delays in progression during their school career. For those who leave school, prospects for employment are difficult, particularly for those resident in rural former bantustan regions. Migration to towns and cities to look for or take up work is a tactic employed by increasing numbers of young people as they approach full adulthood. However, unemployment remains high, and is highest among young women from rural areas.

2.01.2 Sexual behaviour and risk of HIV infection

The economic realities faced by South African young people are challenging. Yet perhaps an even greater threat faces many South Africans during this period of life. Since the early 1990s, levels of HIV infection in South Africa have risen at an alarming rate and by 2001 it was estimated that nearly one quarter of the adult population were infected (32). Just as the adolescent and young adulthood period sees movement away from school and home towards work opportunities, so it sees the formation of early sexual relationships and increasing risk of exposure to HIV/AIDS. Unprotected and risky sexual intercourse remain prevalent, and teenage pregnancy remains high among young women (33, 34). Many are at risk of HIV/AIDS. The following paragraphs will describe levels and determinants of sexual behaviour and HIV infection in South Africa.

While a decade ago there was little literature on sexual behaviour in South Africa, the last 10 years have seen a much greater concentration on this topic in the wake of the HIV/AIDS epidemic. It is often not possible to compare studies across time or place, since a wide variety of indicators have been used in these studies, and the populations they sampled differed in terms of age or other important characteristics. Nevertheless, where possible, these dimensions are explored below.

The timing of sexual debut is influenced both by individual (including biological) and societal determinants. Literature in South Africa has suggested that the age of sexual debut has fallen significantly over the past 20 years (35), although this is suggested largely by increasing rates of teenage pregnancy rather than direct study. It is also clear that young people are getting married later, and thus the period between onset of sexual intercourse and marriage is widening for the majority (36). Traditional norms that regulated youth sexual behaviour, through, for example, peer pressure systems and practices such as virginity testing, have been eroded by colonialism, Christianity and apartheid and the labour migration on which it rested, and have not been replaced by other norms (37). In the present day, young men are seen as in 'need' of sex soon after the physical onset of puberty (38, 39). Gifts may be used by boys to quicken the onset of sexual activity within relationships (40). Yet early sexual activity remains largely hidden from adults. First sex is largely unplanned, rarely protected by contraception and is also often coerced or forced for young women (33, 34). Where contraception is used, condoms appear to be used most often, as opposed to practices such as injectable contraception which require planning, though these become the norm for later sexual activity (33).

Quantitative studies have often included data on sexual debut. Often this is presented in the form of a percentage of individuals ever sexually active. This is difficult to interpret in terms of an average age of sexual debut unless breakdown is given by age (41). In some studies from South Africa data have been presented on average age of sexual debut among populations that are not completely sexually active. This is likely to result in overestimation of the average age of sexual debut. The most appropriate methods for estimating average age of sexual debut in such situations are inspection of the agecumulative distributions of "ever sex", or the use of survival analysis techniques. Neither have been widely applied in South Africa. Nevertheless, data from a wide range of studies put the median age of sexual debut in South Africa somewhere between 16 and 18 years (36, 42-46). Two recent nationally representative studies presented slightly different. The "Nelson Mandela / HSRC" study, concentrating on adults aged 15-49 suggested an average age of debut of 18 years among both sexes (45). Conversely, the "LoveLife" study, which focused only on youth, suggested a lower figure, nearer to 16 years (46). These data would be in line with a general fall in age of sexual debut over the past years, although neither study directly addressed this question. There is little clear pattern on whether age of sexual debut is younger among boys or girls in the South African literature.

Recent literature from South Africa has pointed to high levels of "secondary abstinence" among South African youth (those who are not currently engaging in sexual intercourse after previously being sexually active). This has been suggested by some as a potentially important and neglected component of South African youth's response to HIV/AIDS (35). However, others have suggested that secondary abstinence may simply reflect a lack of opportunity to engage in sex among young people (A.Pettifor, 2004, personal communication). Estimates of the prevalence of the phenomenon include that 7% of previously sexually active women interviewed in the 1998 South African DHS reported no sexual activity for over 24 months (36). In the recent HSRC study, 18% of males aged 15-24 and 14% of females of the same age who had previously been sexually active reported no sexual activity within one year (45), while in the LoveLife study the equivalent figures were 21% and 13% among males and females respectively (46).

Studies have suggested that social norms in South Africa support the notion that men can, and should, have many sexual partners, while women should not (39). Men are described as "needing" sex and being unable to control themselves because of this biological necessity (38). It has also been suggested that having multiple partners has itself formed part of young men's response to the HIV/AIDS epidemic in South Africa (47).

Quantitative data on this subject are particularly difficult to compare across studies due to the heterogeneity in reporting styles. Among various populations, 50% of sexually active men had three or more lifetime sexual partners (48), 6% reported more than eleven partners (42), 47% reported more than seven lifetime partners in 1998, while 34% reported this in 2000 (49), 45% reported more than three partners (46), and 13.5% of ever sexually active males reported more than one partner in the past 12 months (45). Among females, the figures were lower than for males in all studies that compared them, with figures including 7% reporting more than two lifetime partners (48), 8% having five or more lifetime partners (42), 63% reporting more than two lifetime partners in 1998, and 68% in 2000 (49), 31% reporting more than three lifetime partners (46) and 3.9% of previously sexually active women reporting more than one sexual partner in the past 12 months (45).

Age-mixing of populations is essential for the propagation of infection from one age cohort to the next through heterosexual intercourse. Authors have pointed to the potential for significant power imbalances in relationships characterised by wide age differences in South Africa (34, 50). Young women in South Africa may also seek to form relationships with men that are older and / or are able to provide them with resources (ranging from food for the home or school fees, to luxury items such as new trainers or jewellery), and because of the advantages and esteem such relationships bring. The power imbalances inherent in such relationships in South Africa may limit the capacity of women to negotiate the use of contraception, condoms or even when and whether sex happens at all (40).

Surveys from South Africa have generally reported that men are older than their female partners. In KwaZulu Natal in 1999, girls reported that their partners at first sex were an average of four years older than them (33). In another study it was noted that age gaps also varied with the age of the reporting partner. The youngest group reported that 47% of females aged 14-25 years reported partners who were more than five years their senior, while only 14% of young males reported partners who were more than five years younger (42, 43, 49). A similar pattern was seen in the LoveLife survey, where young males reported partners who were on average one year younger than themselves, while females of the same age group reported that their partners were on average some four years older than them (46).

Resource exchange is particularly complicated to address in standardized questionnaires since the meaning of such exchanges is very subjective. In KwaZulu Natal, 4% of males replied that they had "received money in exchange for sex", while 4% reported that they had "given" such resources.

Among females, the proportion reporting receiving resources was higher at 11%, while 3% reported they had given resources (33).

The majority of youth in South Africa are aware that HIV is present in the country and know that condoms can prevent HIV infection. However, this knowledge co-exists alongside low perception of personal risk in the general population and a number of other barriers to condom use. Condoms are associated widely with decreased sexual pleasure, particularly for men (51). Further, discussion of condom use is often associated with either accusation or admission of infidelity (52), since their use is associated only with casual partners, and not with loving relationships (51). This means that raising condom use within a partnership can be dangerous, particularly for young women in a society where violence within relationships is common. Males expect condom use to be dropped after some time with partners who they do not perceive to be at risk of HIV, and use gifts to signify love for their partners in order to progress to this phase quickly (40). Male peers may laugh and joke at others who report using condoms (51). In addition to these barriers, rumours persist in some rural areas that condoms themselves can lead to HIV infection (53). Some authors have pointed to the importance of proving fertility as a barrier to contraceptive use (54), although this has been challenged in the South African context (34). Condoms are largely seen solely as a means of HIV prevention, and are not widely associated with their contraceptive use by young people (39). Finally, even when young people wish to use condoms they may not be available (51, 55).

With such an array of barriers to condom use, it is unsurprising that reported condom use is far from universal. Nevertheless, what also seems clear is that condom use has become much more widespread in recent years from a time when condom use was rarely, if ever, reported by young people. Ever having used a condom was reported by 63% of males in both 1998 and 2000, and by 50% of females in 1998 and 59% in 2000, after two years of an intervention programme (49). A more commonly reported marker has been that of condom use at last sex with a non-spousal partner. In a national study, 57% of sexually active males and 46% of females aged 15-24 reported condom use at last sex, and these figures were higher than among the elder groups in the survey (45). Very similar figures were reported by another more recent national survey, with 57% of males and 48% of females reporting this (46). Lower reported rates by females may reflect that they have older male partners who are less likely to have changed their behaviour in response to HIV/AIDS. Finally, some authors have measured consistency of condom use. In one study, the numbers of males aged 15-24 with a regular partner who reported always using a condom increased from 18.9% in 1998 to 44.4% in 2000, while among females, there was a smaller increase over the same period (from 22.9% to 29.2%). A national survey conducted in 2002 reported 33% of young people of both sexes reported that they always used a condom with their most recent partner (46). Condom use has been shown to be more common among casual and non-marital partnerships than within spousal partnerships (56).

Rape and sexual coercion remain at alarmingly high levels in South Africa (57-59), although the reasons for this are unclear. Rape has been described as both a manifestation and an assertion of male dominance over women, and as an extreme manifestation of perceived male sexual entitlement in South Africa (58). Male violence over women is accepted within South African society and is even respected by male peers. Others have commented that the historical legacy of apartheid propagates high levels of sexual violence and other forms of gender-based violence in South Africa (60). The entrenched low economic status of women puts them at risk of sexual violence, while alcohol abuse has also been associated with higher levels of violence (61). A recent national study confirmed a very high level of misconceptions about sexual violence among school students (60). Coercive or forced sex may also be an independent risk factor for HIV infection for women since it is more likely to result in vaginal tearing which may facilitate transmission of the virus (62).

Empirical data on rape and sexual coercion are difficult to collect. Nevertheless, there have been some attempts to do so in South Africa. The South African Demographic and Health Survey found that 7% of women aged 15-49 reported non-consensual sex during their lifetime (36), and 1.6% of women reported being raped before 15 years of age (58). In another study, 4.5%-7.2% of women across three provinces reported lifetime experience of rape (57). More recently, a study among antenatal clinic attending women in Soweto suggested that 20.1% had experienced sexual violence at some time, most often in conjunction with physical violence, and that HIV infection rates were higher among the group experiencing either form of violence even after adjustment for associated HIV risk factors (63). Among school students aged 11-19 years, being forced to have sex was reported by more males than females at the youngest ages (10% vs. 6%), but by fewer males than females at older ages (12% vs. 15%) (60). The most recent data among South African youth come from a national survey of 15-24 year olds, in which 2% of males and 10% of females reported ever physically being forced to have sex (46).

It is widely accepted that heterosexual sex remains the most important source of HIV transmission in South Africa, though this has been disputed (64). This is by far the most important source of transmission among adolescents and young adults (65) and the epidemic spread of HIV has been to a great extent driven by the characteristics of sexual behaviour practice described above.

Additionally, a number of co-factors for HIV transmission have been described that may have influenced the spread of HIV within South Africa. Sexually transmitted disease levels are high, and in particular HSV-2 infection has been found to be a strong risk factor for HIV infection in South Africa in line with findings in other countries (42, 66). Male circumcision has also been described as an important risk factor for HIV infection in other settings (67), and may be an effective intervention for HIV prevention in South Africa where circumcision is common among some ethnic groups but not among others (68-70). Practices that may increase risk of HIV transmission such as dry sex, female circumcision and vaginal douching have not been widely reported in South Africa.

The first cases of AIDS in South Africa were detected among homosexual men in Cape Town in 1982, and were the result of infection with HIV clade B. In the more than 20 years since then, the homosexual epidemic has been dwarfed by an explosive spread of clade C HIV infection among the general population, with heterosexual and perinatal being by far the most important modes of transmission. The most rapid rise in HIV prevalence in South Africa was seen between 1993 and 1998. This period is later than the spread of HIV in countries to the north of South Africa. UNAIDS estimates show that HIV spread was most rapid in East Africa during the 1980s. For example, the prevalence of HIV in Uganda was 20% in 1987 when UNAIDS record keeping began. National HIV prevalence in Zimbabwe first exceeded 10% in 1989, yet national prevalence first exceeded 5% in 1990 in Botswana, 1994 in South Africa and 1996 in Mozambique.

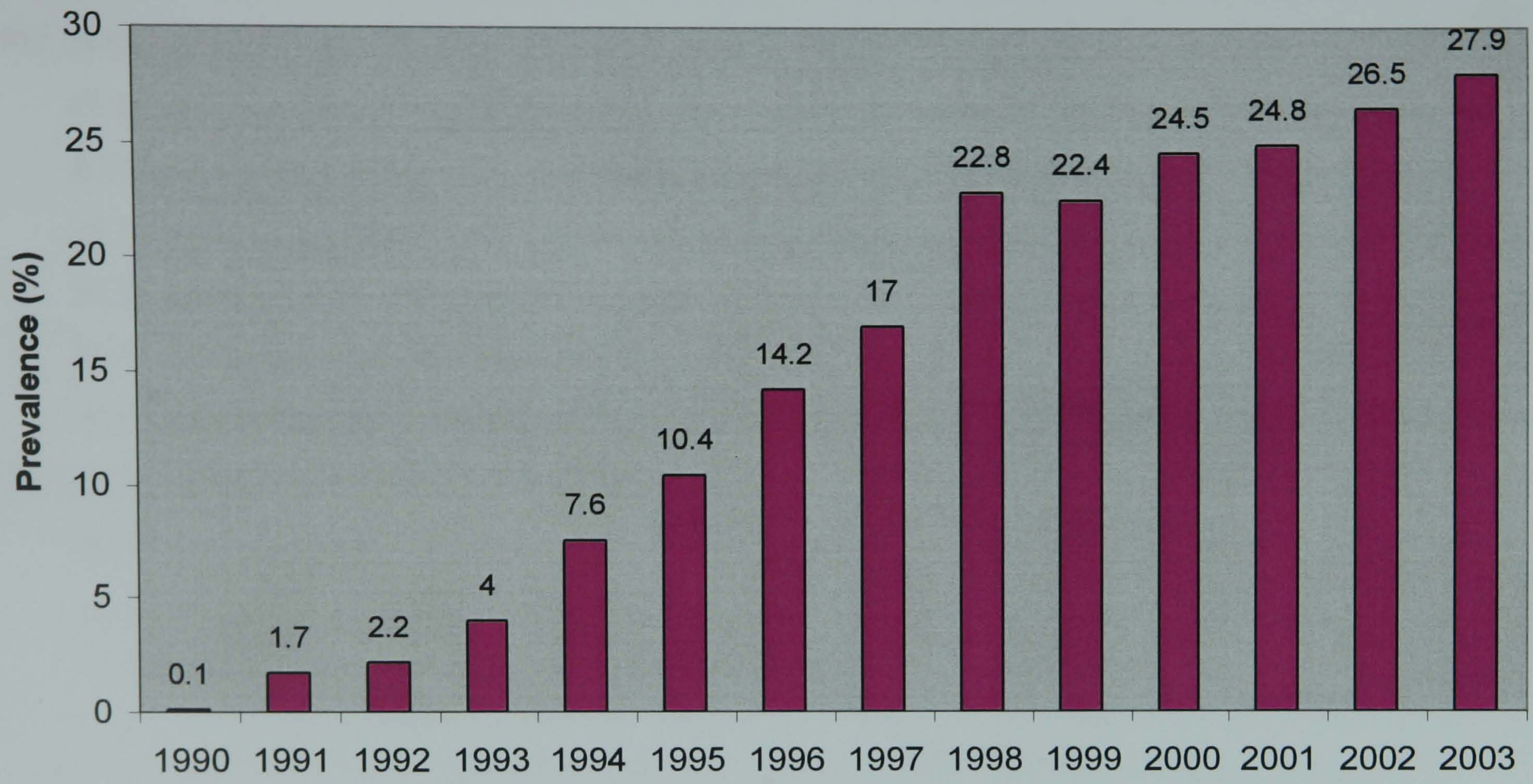
Estimates of the national prevalence of HIV infection in South Africa among the general population come from a now well-established annual antenatal clinic surveillance system that has been operating since 1990. These data are effective in tracking the population level of infection over time. However, biases inherent in antenatal clinic data may have particular relevance to young people. In particular, since a relatively high proportion of the youngest women may continue to abstain from sexual intercourse, antenatal clinic data may tend to overestimate HIV prevalence in this group. Antenatal data have in recent years been supplemented by two nationally representative serosurveys conducted in South Africa.

Figure 2.1 shows the prevalence of HIV infection among antenatal clinic attending women in South Africa from 1990 – 2003. A pattern of explosive spread can be seen over the early years of the 1990s, with continued slower increases in prevalence beyond 2000. In 2001 the national prevalence of HIV infection was 24.8%. This figure masks considerable regional differences.

Prevalence was highest in the province of KwaZulu Natal, with one third of all women tested found to be HIV positive, while estimates for Mpumalanga and Limpopo Province (where the current study was conducted) were 29.2% and 14.5% respectively. The prevalence of HIV infection also varies considerably with age. Data from a nationally representative study conducted in 2001 show that prevalent infections peak at 25-29 years for women and 30-34 years for men (see Figure 2.2).

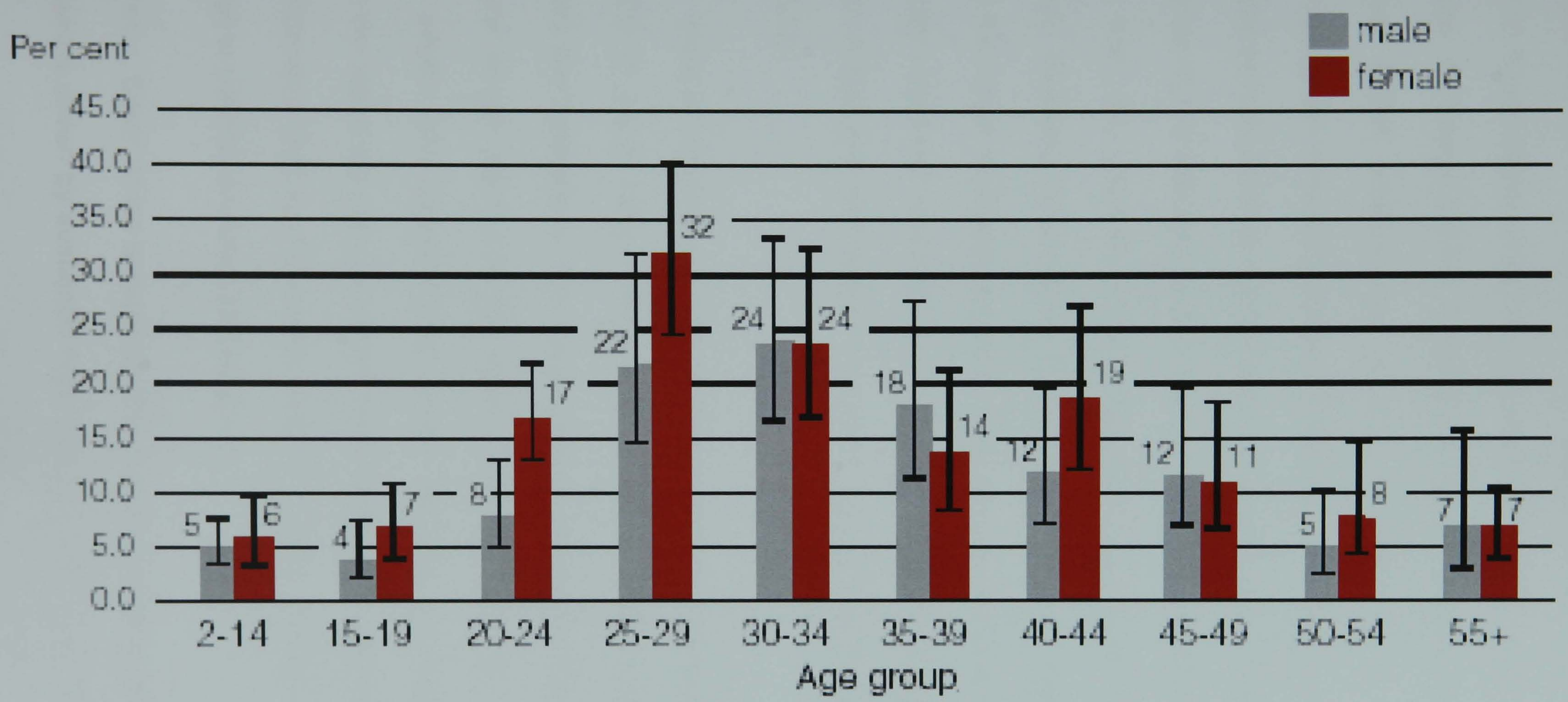
Two recent national surveys have added to understanding of the pattern of HIV infection in South Africa, and in particular provided more appropriate information for understanding the burden of infection among young people and males. The Nelson Mandela / HSRC study conducted in 2001 found a national prevalence of HIV infection of 6.1% among males and 12.0% among females aged 15-24 years (45). This suggested a similar pattern of infection to that seen in other studies in Africa. Females at young ages were significantly more likely to be infected than males (see Figure 2.2). This study confirmed that among this age group antenatal clinic data overestimated the prevalence of HIV infection, primarily because a large number of individuals aged 15-24 years have not yet become sexually active, and this group are not represented in the ANC sample. The national LoveLife survey of 15-24 year olds conducted in the same year found similar results, with 4.8% of males and 15.5% of females aged 15-24 being infected (46). Both studies also provided provincial level data of relevance to this thesis. In the HSRC study, HIV prevalence was 5.6% in Limpopo province and 11.7% in Mpumalanga, while in the LoveLife study HIV prevalence was 6.7% in Limpopo and 11.7% in Mpumalanga.

Figure 2.1: Prevalence of HIV infection in South Africa among women attending antenatal clinics, 1990 – 2003



Reproduced from National HIV and syphilis ante-natal seroprevalence survey in South Africa 2003 http://www.doh.gov.za/docs/reports/2003/hiv/p1-23.pdf.

Figure 2.2: Prevalence of HIV infection by age in South Africa among men and women in 2001



Reproduced from Nelson Mandela / HSRC Study of AIDS 2002 (http://www.hsrcpublishers.co.za/e-books/HSRC_HIV_survey.pdf)

Directly measured data on incidence of infection are rare in sub-Saharan Africa, and there are no reports of such measures among general population groups in South Africa. However, attempts have been made to estimate HIV incidence in South African populations using model projections and a "detuned" ELISA methodology. Among a population of antenatal clinic attending women in KwaZulu Natal in 1999, the incidence of HIV infection was estimated by both of these techniques (71). Among women aged 15-19 years, prevalence of HIV infection was 24.5%, while the estimates of incidence were 18.9% from the detuned assay approach and 15.4% from the model approach. Between 20-24 years the respective figures were 40.0%, 23.6% and 26.6%. Peak incidence was estimated to occur around 22 years of age. This study confirmed very high levels of HIV incidence among this population in South Africa, with the highest incidence occurring in the young adulthood period (72).

Finally, there is increasing evidence of high levels of mortality as a result of HIV/AIDS in South Africa. Despite limitations in vital registration systems estimates are available on this issue. These suggest significant rises in both male and female adult mortality since 1985. In 1999/2000, male mortality among adults was approximately double the level in 1985, while among women the figure was almost a three fold increase. In the same year, it has been estimated that approximately 40% of deaths among 15-49 year olds were due to HIV/AIDS related illnesses (73).

In summary, South Africa is experiencing a significant generalised epidemic of HIV infection driven by sustained high levels of risky sexual behaviours such as multiple partners for men, age asymmetries in sexual relationships and lack of condom use. While levels of HIV infection are low among children, they rise rapidly during the late teens and early twenties, particularly among young women, suggesting that incidence of infection is high among this group. Women in this age range are at significantly higher risk of HIV infection than men. Factors that increase the risk of transmission within unprotected sexual acts, such as the presence of STIs, lack of male circumcision and high levels of violent sex, are also common in South Africa, further facilitating spread.

2.02 Theoretical basis for a social epidemiology of HIV infection

The primary aim of this thesis was to describe the social epidemiology of sexual behaviour and HIV infection among unmarried young people from a rural South African population. This section will describe theoretical and conceptual literature underpinning this empirical aim. It will describe theories that help to identify and conceptually link relevant socioeconomic exposure variables, before moving on to theories of the proximate determinants of sexual behaviour and the epidemiology of HIV infection. This section will end with a description of theoretical literature that links these exposures and outcomes.

2.02.1 Conceptualising and measuring socioeconomic status

(a) Theory

A number of disciplines, notably sociology and economics, have sought to explain social structure and the generation of well-being in society, while more recently social epidemiologists have sought to use relevant classifications from these fields to inform their study of the relationship between social status and health. A brief review of important concepts from these fields is useful to help identify and organise relevant socioeconomic exposures in a manner that will inform a better understanding of the social epidemiology of the risk of HIV infection.

Sociologists have long debated the processes by which social classes are formed and can be identified. Broadly, Marxist sociologists describe society in terms of eternal conflict between social classes defined by their position in relation to the means of production (thus, capitalists and labourers). Weberian sociologists suggest greater stratification, in which social classes might be identified along at least three potentially overlapping axes, those of ownership, status and power (74). Nancy Krieger and colleagues have sought to operationalise these concepts for use in epidemiological study in modern economies (75). They identify two distinct aspects of social stratification; socioeconomic position and class. Socioeconomic position is an ordered continuous variable describing the place occupied in distributions of, for example, income, wealth or education. It might be considered on either an absolute or relative scale. Conversely, class is a nominal variable, defined on the basis of "social groups arising from interdependent economic relationships", and is often considered in terms of occupational or educational classes.

It has been noted that relevant social class variables for children and adolescents remain unclear, since the level at which status should be measured is not immediately obvious (74). Krieger et al recognise that position or class variables might each be captured at individual, household and community levels. The household may be a particularly important level of measurement for understanding the welfare of the young (76). Economists and sociologists each recognise the potential importance of the household as a unit that is exposed to, buffers individuals from and creates the social and economic conditions in which people live. In Southern Africa, the concept of household is linked to family, kinship and co-residence yet is not synonymous with any of these (77). Notably, the unique 20th century history of South Africa has led to a situation where linkage between rural households and migrants who are usually considered members of these households is of particular importance (12, 78).

Modern economic theories identify that households seek to maximise their utility through adopting a portfolio of activities. Frameworks developed for use in low-resource settings identify multiple sources of capital on which households and their members draw to maximise well-being (79, 80). Depending on the economic environment, households might sell labour or produce goods for sale in the marketplace. As alluded to above, migration to seek or take up work is a particularly important livelihood strategy in many settings (81). Long term strategies are of particular relevance when considering young people. In particular, investing in the continued education of young people can be seen as a long term investment strategy for the household. The capacity of the household to make this investment may, however, be dependent on more immediate needs and the costs associated with continued education.

Drawing on these frameworks, relevant variables were sought that would help improve an understanding of the social patterning of risk of HIV infection. Previous social epidemiology in South Africa has identified race, socioeconomic position and migration as key variables (82). The population in the current study was expected to be homogenous with regard standard classifications of race so this was not a suitable variable for the current study. However, both socioeconomic position and migration were deemed highly relevant. Further, among young people it was also felt that current attendance at school or college would be an important factor, with some young people remaining in education and others leaving education.

Thus, as will be seen in the conceptual framework depicted at the end of this chapter (Figure 2.4), three socioeconomic variables were organised in a manner informed by these frameworks. The first variable was relative household wealth, which was defined in terms of the relative level of welfare (or standard of living) experienced by household members that was the product of socioeconomic activities. It was assumed that the welfare of young people would be primarily determined by the overall wealth level of the household rather than by their individual activities. However, it was hypothesised that the level of household wealth might affect the socioeconomic role that young people were currently engaged in. Thus, wealth might affect whether young people were attending school or college or were no longer doing so, and this was the second variable considered for examination in this thesis. Further, it was hypothesised that some young people would be temporarily migrant while others would be at home, and thus that this should be the third exposure variable. Migration may be for the purposes of schooling, or may be for the purpose of work seeking or taking up employment and might also have been affected by household wealth and / or by school attendance.

In developing this framework it was recognised that the role adopted by young people may itself impact the wealth of the household. This might happen for example, through a young person contributing to the household income if they successfully take up employment. However, it was assumed that in most cases the socioeconomic roles adopted by young people are not the primary determinant of household wealth. As will be seen, the majority of young people who had left school had not successfully found stable employment, but were seeking employment and or had found occasional low-paid work.

The pathways through which these factors might influence risk of HIV infection will be described in greater detail later. However, it should be noted here that it was felt that household wealth might most directly affect risk of HIV infection through an impact on self confidence or perceived social status, or by directly affecting the capacity to provide, or need to receive, resources within sexual relationships. Wealth might also have had an impact on HIV risk characteristics through pathways in which it proximately affected attendance at school or migration status. These variables were assumed to have the potential to most directly influence risk of HIV infection through affecting exposure to HIV prevention messages and sexual network structures.

(b) Measurement

Finally in this section it is necessary to briefly discuss measurement issues. While data on socioeconomic role (attendance at school / temporary migration) seem relatively simple to capture, a wide array of potential choices were available for the measurement of household wealth. The choice of marker must be made on both conceptual and methodological grounds. This study required the development of a marker of socioeconomic position that was suitable for use among a population of young people and allowed the study participants to be ranked. Throughout this thesis the term "household wealth" will be used to describe the underlying concept these markers sought to measure. This was intended as a broader term than household income, standing wealth and human and social capital; indeed its use is intended to reflect the combined influence of such socioeconomic assets in generating welfare for household members (80). Those at the top of the ranking would be expected to have the highest welfare levels, and would be referred to as the "wealthiest" group, while those at the bottom of the ranking would be expected to experience the lowest welfare levels (the "poorest"). Methodological issues in developing such a scale are discussed below, and formed an important component of this thesis, seeking to improve on previous studies in this field that have rarely described a theoretical basis for the markers developed (83). It should be noted that in economics "wealth" can have a more precise meaning, that of the monetary value of owned assets (which can be contrasted with income), but this more precise meaning is not intended here. Finally, as will be seen, the measures developed in this thesis did not use monetary information, and so measured only relative wealth within the population as opposed to absolute wealth measures that could be compared beyond the population borders, although aspects of the absolute welfare level experienced by those at the top and bottom of the scale will be described.

Economists have generally adopted money-metric measures to assess household welfare (84). Under this model, the consumption of consumer goods may be regarded as the most appropriate measure of current standard of living, but is conceptually difficult to measure. Consequently, surveys more often measure expenditure on consumer goods as a proxy for their consumption. Alternatively, they may measure income since this provides a good measure of "opportunity to consume". The poor are identified through the application of a poverty line to the collected data (85).

There are both conceptual and methodological challenges to this approach (86). It is increasingly accepted that standard income and consumption measures reflect a relatively static conception of well-being. Additionally it has been suggested that monetary measures fail to capture other important aspects of welfare such as community resources, social relations, culture, personal security and the environment (86). Methodological limitations include that sensitive data on income may be subject to under-reporting, while expenditure information is susceptible to recall bias. It may be difficult to estimate the value of home-produced goods and the use-value of services such as water, electricity, education or health services. There also remains considerable debate on the application of equivalence scales to account for household size and composition. Finally, in addition to these driving forces, and perhaps most importantly, the collection and analysis of income and expenditure data are complex and labour and time intensive. Health researchers rarely have the time, resources or expertise necessary to collect such data. In response to this constellation of forces, there is considerable interest in alternative techniques for the estimation of household wealth.

A variety of rapid household wealth measures that use survey data have been developed, including shortened income or expenditure questionnaires, evaluations of housing quality or other visual indicators of well-being, measurements of educational or nutritional status and many others (87). In recent years there has been particular interest in using statistical techniques to combine data on ownership of assets, access to services and/or other markers of welfare. The aggregation and weighting of such data might be achieved in a number of ways – including simple addition, addition with weighting based on local consultation or estimated prices, or through the application of statistical procedures. A number of authors have recently used principal components or factor analyses to generate a household wealth index. This approach was pioneered using DHS data and is gaining popularity (88-93).

Health researchers might also look to broaden the range of techniques available to them. Participatory rural appraisal (PRA) represents an approach to information gathering quite different from survey methodology (94). PRA is a family of exercises performed with groups of individuals to promote analysis of problems. The techniques seek to be rapid and accurate, but also to involve communities in the process of data collection and analysis. There is a wide variety of tools, but each tends to emphasise local knowledge and promote participation and teamwork. One such technique is Participatory Wealth Ranking (PWR), in which community members themselves rank the relative wealth of households in their community (95). These tools have been widely adopted in development practice, but have rarely been used in health studies. While a number of studies have suggested that PWR can provide valid information on relative wealth (96-98), there remains much to learn about the technique.

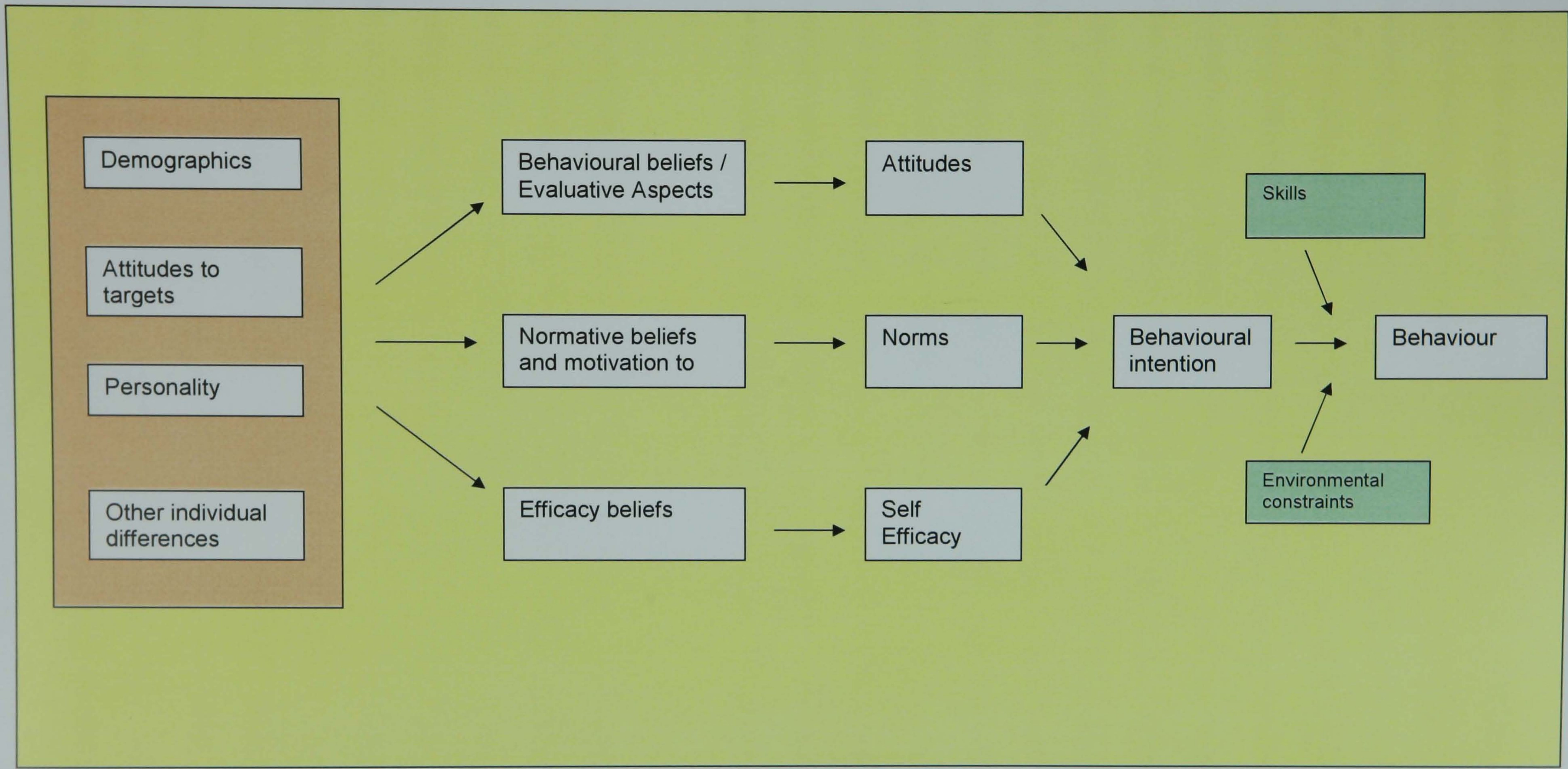
As will be seen, this study sought to develop the most appropriate marker of household wealth possible from the available data. Three measures were developed, encompassing data from both survey and PWR methods. The method that was judged the most appropriate on the available information was selected for use in the main study.

2.02.2 Theories of sexual behaviour and behaviour change

A range of theories describe motivations for the behaviour of individuals and processes of behaviour change. These can be applied to promote understanding of sexual behaviour and how this changes in response to a health threat such as the acquisition of HIV infection.

Theories such as the theory of reasoned action and theory of planned behaviour (99), the health belief model (100) and social cognitive theory (101) have some properties in common. In particular, they focus on individual cognitive and psychosocial determinants of health behaviour. An integration of these theories is shown in Figure 2.3 (102). Under this model, intention to perform a particular behaviour (for example, to use a condom with a particular partner) is a strong predictor of adoption of the behaviour. The likelihood of the behaviour being performed is increased if the individual additionally possesses the skills to perform the behaviour, and there are no environmental constraints to it being performed. Developing an intention to perform the behaviour is itself the product of the individual's attitudes about the behaviour, their perception of social norms relating to the behaviour, and their belief in their ability to perform the behaviour (self-efficacy). Each of these factors is in turn a function of a set of beliefs about the consequences of the behaviour, the behaviours and attitudes of peers and the barriers to performing the behaviour. Finally, these beliefs are themselves the product of a wide range of socio-demographic and other factors. These theories have highlighted the importance of specificity when discussing behaviours, since, for example, condom use with one partner may have different implications than condom use with another. In addition, such theories are particularly useful when combined with other individual theories such as the stages of change model (103), which seeks to target behaviour change messaging by recognising a spectrum of positions individuals may be in regarding a specific behaviour. Specifically, this model identifies individuals in the following classes; pre-contemplation, contemplation, preparation, action, maintenance and relapse. Used appropriately, such theories have been shown to be highly effective in predicting behaviour and have been used to develop and evaluate targeted interventions in some settings. These models have also been adapted specifically for AIDS programming, for example in the AIDS risk reduction model (104).

Figure 2.3: An integrative model for sexual behaviour, reproduced from Fishbein (2000)



Some authors have suggested that these theories share a common limitation in their focus on individual level determinants of behaviour (105). A number of other behavioural theories explicitly create frameworks highlighting that behaviour and behaviour change are not solely individually driven. Diffusion of innovation theory describes the processes by which innovative behaviours are adopted by communities. It suggests new behaviours are first adopted by a small number of "innovators" (2-3% of the population). Following this, "early adapters" (often more educated or wealthier individuals comprising 10-15% of the population) adopt the innovation and lead both the "early majority" (30-35%) and "late majority" (30-35%) over time into adopting the behaviour. A relatively conservative group (the "laggards") adopt new trends only after some time or never adopt them. Some HIV prevention interventions, for example peer education methods, aim to speed up the process by which wide scale community adoption of a safe behavioural practice (such as condom use) occurs.

Other authors emphasise that sexual behaviour is not individually determined. Social network theories point specifically to the fact that by definition sexual relationships involve more than one person, pointing to another layer of factors (including gender and power) that can affect behaviours within sexual partnerships. Campbell has emphasised the role of social capital in fostering safe sexual behaviours (106, 107). Under this theory, individuals are seen as collectively negotiating social identities and establishing group norms about specific behaviours. In order to foster effective responses to HIV, such as the adoption of safe sexual behaviours, individuals and groups must be empowered to develop a critical consciousness relating to the issues affecting them. To do so it is necessary to access stores of both bonding social capital, which is necessary for the formation of strong, supportive group identities, and bridging social capital which links groups together, in particular linking the relatively underpowered with those with more influence over external factors.

In summary, there are a range of theories that describe the behaviour of individuals and social groups. Such theories provide insight into the generation of sexual behaviour and HIV risk, but usually fall short of explicitly describing the mechanisms by which socioeconomic forces might influence HIV transmission.

2.02.3 Epidemiological theories of HIV transmission

Epidemiological theory has been used to describe how patterns of sexual behaviour and the distribution of co-factors can lead to patterns of HIV infection within populations. It is widely (65), although not universally (108), accepted that heterosexual transmission remains the most important route of HIV transmission in sub-Saharan Africa and this is particularly true among young adult populations.

The following equation can be used to describe the heterosexual transmission potential of HIV within populations (109).

$$R = b.\ c.\ D$$

Where R = the reproductive number for HIV (a measure of the transmission potential of the virus in a given population), b = the probability of sexual transmission from one infected person to an uninfected partner, c = the average rate of new partner acquisition, and D = the average period of infectiousness. By breaking down this equation, a number of important empirically observed determinants of HIV transmission can be understood.

The rate of spread of HIV is firstly determined by 'b' the probability of transmission. This is influenced by both sexual behaviour and other factors. Sexual behaviour factors include the probability that the other partner is infected with HIV. If neither partner is infected, or if both partners are infected, no new seroconversion will occur. A number of sexual behaviours may influence the likelihood that a partner is newly infected during sexual intercourse. Firstly, contact with individuals from populations with a higher background level of infection are more likely to result in transmission. Such groups include commercial sex workers (who are usually female), long distance truck drivers (usually male) (110) and individuals from urban (compared to rural) populations (45). Secondly, older individuals have had a longer period in which to become infected. Consequently, significant age mixing in populations is likely to increase transmission (111). In particular, in many settings where men have younger female partners, this may pose a particular risk for females (112). Thirdly, some sexual behaviour practices have been suggested to influence risk of transmission, including vaginal douching and drying (113). Fourthly, correct use of condoms reduces the risk of transmission to almost zero within sexual partnerships. Condom use has risen in Uganda where HIV prevalence has fallen steeply in the past few years (114). Finally, it has been suggested that females may be at greater risk of infection from males than vice versa; younger females may be more susceptible than older females; and that women experiencing violent or coercive sex may be at particularly high risk of transmission (62).

Sexual-behaviour co-factors may also affect the risk of transmission. It has been found that the presence of both ulcerative and, to a lesser extent, non-ulcerative STIs in either partner increases the risk of HIV transmission (115), although the role of STI prevention programmes in HIV prevention in different epidemic settings remains unclear (116-118). A large number of studies have suggested that male circumcision is protective of HIV transmission (67, 119), although this remains to be confirmed in randomised controlled trials of circumcision. Further, high viral load, particularly associated with the early and late stages of infection, is thought to increase the probability of transmission.

Regarding behaviour, the rate of partner change ('c' in the equation above) has been shown to be associated with epidemic spread in a number of settings. Modelling studies have additionally suggested that concurrency of partnerships, as opposed to serial monogamy, is particularly conducive to rapid spread of HIV (120-122). The age of sexual debut is also associated with total numbers of sexual partners, while younger age at sexual debut is linked to HIV risk through both a biological mechanism and through its affect on level of sexual exposure to HIV.

For infections such as HIV that confer lifelong infectiousness, the major determinant of duration of infectiousness ('D') is age. Thus the length of time an individual is infectious might be increased by antiretroviral therapy programmes that are successful in extending the life of individuals who are HIV positive. However, this might also be weighed against the fact that effective treatment programmes are likely to lower viral load, thus decreasing another important parameter in the equation.

Distal factors, including economic forces, have the potential to influence the transmission of HIV if they influence any of the factors described above (b, c or D). The relationship between these different components means that distal forces might be likely to act to increase some components and decrease other. The final section below discusses theories that explicitly relate economic forces to specific components of the equation above are discussed.

2.02.4 Theories on the social determinants of HIV risk

Economic, behavioural and epidemiological theories so far presented in this section have been primarily concerned with understanding the proximal determinants of, in turn, social class, sexual behaviour and HIV epidemiology. These theories do not exclude the possibility for there to be more wide ranging effects or determinants of the systems they describe, but these are not the focus of their concern. There has however been considerable literature describing in more detail the pathways that might link economic factors and sexual behaviour or HIV infection.

Some such theories have explicitly targeted an understanding of societal level determinants of epidemic HIV spread. Thus, HIV epidemics may have different courses within societies categorized by risk environments that are determined by their wealth and social cohesion (123). Societies that are wealthier, have more equitable income distribution and are more gender egalitarian, and those that have a higher human development index, tend to have less severe HIV epidemics (124, 125). These theories do not explicitly describe how socioeconomic forces shape the risk of individuals within a given society.

The literature on structural barriers and facilitators to HIV prevention considers both societal and individual determinants of vulnerability (126, 127). Internationally, three major forces are highlighted; poverty, mobility and gender inequality (128). Other structural forces are also important, in particular political will. It has also been suggested that the relevance of such structural forces might be expected to change over time (129). The sections below will outline some of the pathways through which wealth, education, workforce participation and migration might influence the sexual behaviour and risk of HIV infection of young people in South Africa. Where possible, gender differences in the nature or importance of these pathways will be highlighted. It should be noted that such pathways do not generally challenge the importance of the models of sexual behaviour and HIV epidemiology described in the previous sections. Rather, they rely on such theories as the building blocks through which the pathways operate. For example, demographic and epidemiological approaches have been combined in a proximate-determinants framework applied to HIV/AIDS transmission (130). In South Africa, a recent review provided a framework for organizing the relationship between sexual behaviour and factors at the personal level, and the proximal (interpersonal and physical / organizational environment) and distal context (culture and structural factors) in which these behaviours occur (131).

Social status can affect health through psychosocial pathways. The poorest members of society may have relatively low self esteem, self efficacy and a negative outlook on the future (132). They may be less likely to adopt health protecting behaviours such as having fewer sexual partners or using condoms. This will be particularly true when such behaviours are deemed undesirable by peers for a variety of additional socio-cultural reasons. Put simply, among many competing priorities there may be little incentive for the poorest members of a society to avoid HIV infection.

In many African settings it remains socially acceptable for men to have many sexual partners, but unacceptable for women to do so (39, 50). Within these same settings, women are often more likely to leave school early, be unemployed and to earn low wages when they are employed (133). Consequently, men in employment or from wealthier households may be more attractive to women and their families as marriage partners or boyfriends, and may have greater numbers of sexual partners. Such men may also have a greater opportunity to engage with commercial sex workers who are at high risk of infection. This situation puts both men and women at high risk of heterosexual infection. Women, unable to access labour markets and with low levels of educational attainment, may increasingly rely on sexual relationships as a way to access resources, and young women remain the most disenfranchised and at risk population group in many settings (133, 134). Some may enter formal commercial sex work. More often women enter sexual relationships characterised by significant power imbalance. The formation of relationships characterized by wide age gaps is seen as culturally normal, but some have suggested it may also be driven by economic forces. Older males might be seen as better options as boyfriends and husbands because they are perceived as more able to provide resources to support the woman and her children (40). In turn, age asymmetry and economic dependence are linked to decreases in the power of women to negotiate within sexual relationships. Gifts and money are an intrinsic component of many sexual relationships, with gift giving being seen as the norm by both males and females (40). Men often dictate the timing and characteristics of sexual intercourse. Since men often perceive condom use negatively for a wide variety of reasons (51), their wishes not to use condoms may prevail. Greater personal income for men, particularly those who are working, may also afford them the opportunity to negotiate sex without condoms (40). Additionally, commercial sex workers may charge a higher price for "skin on skin" sex, making this only accessible to men with a higher income. The poorest women may also be those at the greatest risk of sexual violence.

In many settings, knowledge of the protective effect of condom use is high, yet there remain significant barriers to condom use including negative peer influence, distrust, dislike and lack of availability (51). Attendance at school may be linked to higher levels of contact with and understanding of HIV prevention messages, and may also foster group norms that embrace positive values with regard to condom use (135). Schools also offer the opportunity to join groups that provide access to both bridging and bonding social capital, which might also be important in fostering positive responses to HIV prevention messaging (136). Education might also confer a sense of optimism about the future, and foster critical engagement with health promotion literature (137), leading to the adoption of safer sexual behaviour traits. Finally, in many countries of sub-Saharan Africa first marriage is increasingly occurring at later ages particularly among the most educated, while there has been little change over time, or perhaps a reduction, in average age at first sex (138). Consequently, the number of years between onset of sexual intercourse and marriage is increasing for all young people, but may be increasing most rapidly among the most educated groups.

Accessing appropriate treatment for STIs may also be influenced by wealth and education. Consequently, those of lower socioeconomic status may suffer higher burdens of STI, and consequently higher levels of HIV infection.

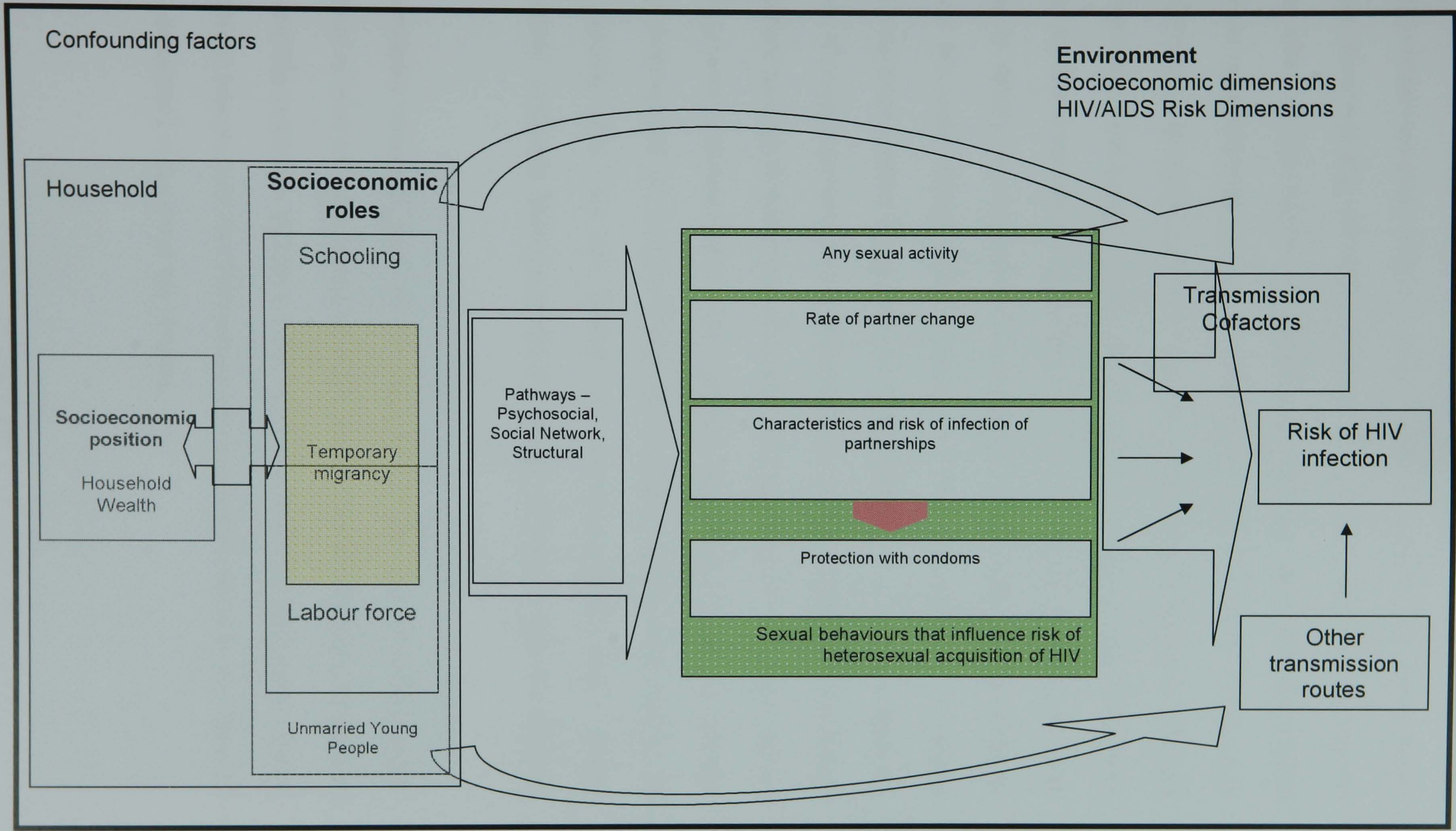
In settings where labour migration is common, such as in South Africa, there are further drivers for high numbers of sexual partners. Men who migrate, may have partners both at the rural home and at the destination of migration, including separate families and girlfriends (139). Migration might also form part of the process and construction of masculinity and part of a transition to manhood for young men (122, 140). Migrant men may have more opportunities to engage in casual relationships or commercial sex work for reasons including loneliness and isolation (141, 142). The prevalence of HIV infection among commercial sex workers is often high, and consequently, contact with these groups is linked to a higher risk of HIV infection for clients. In addition, male migrant workers with greater income may have more opportunity for contact with commercial sex workers (106). Even in the absence of contact with sex workers, across sub-Saharan Africa urban areas have recorded higher levels of infection than rural areas (45). Consequently, migration to urban areas from rural areas, and the formation of sexual relationships in these higher HIV prevalence environments, can carry with it a higher risk of infection for both men and women, even if absolute levels of sexual activity are comparable to non-migrants. However, migration from rural to urban areas may also be associated with increased contact with HIV prevention messaging, through workplace programmes that have become widespread in South Africa, and increased accessibility of condoms in migration destinations. Women who remain at home while their husbands migrate may themselves also be likely to engage in more sexual partnerships, and may form relationships with individuals who themselves have migrated into an area (143, 144). In addition, female migration is increasing in frequency, and women migrants may be expected to provide sex as a form of payment for travel, accommodation or other resources (145). Finally, long term separation may be a barrier to male and female partners involved in sexual relationships discussing issues such as condom use and other sensitive matters such as HIV testing.

2.03 Conceptual Framework

On the basis of the literature reviewed above, a conceptual framework was developed to guide the investigation of how socioeconomic factors relevant to rural South Africa might influence risk of HIV infection. This framework is depicted in Figure 2.4.

The conceptual framework situates *unmarried young people* of either sex within *households* that are themselves situated within a broader *environment*. Two dimensions of this *environment* are of particular importance – the *socioeconomic* and *HIV/AIDS risk* dimensions. The socioeconomic dimensions include the available opportunities for generating household wealth – including investing in education, engaging in agriculture, migration, employment, self-employment or accessing state claiming systems (such as pensions). The environment is also characterised by factors that create differentiation in access to these things including age, gender and mobility. The HIV/AIDS risk dimensions of the environment include levels of infection, accessibility of HIV prevention messages, prevailing behavioural norms, levels of behaviours that influence risk of infection and STI treatment opportunities, as well as the factors that describe differentiation in these, again including age, gender and mobility.

Figure 2.4: Schematic diagram depicting key components of the conceptual framework



Within this *environment,* economic theory suggests that households will attempt to maximise their welfare through a portfolio of activities. In this thesis it will be assumed that adolescent welfare is predominantly affected by the welfare level of the whole household. As part of this portfolio, young household members may continue in education or may leave school and enter the workforce. For either purpose, young people may temporarily migrate away from the rural home or may remain at the rural homestead. Migration may be for the purposes of schooling, or may be for the purpose of work seeking or taking up employment. Household wealth has the potential to influence the roles young people engage in. The role adopted by young people may itself impact the wealth of the household. This might happen for example, through a young person contributing to the household income if they successfully take up employment. However, it is assumed in this thesis that in most cases the socioeconomic roles adopted by young people are not the primary determinant of household wealth. As will be seen, the majority of young people who had left school had not successfully found stable employment, but were seeking employment and finding occasional low-pad work.

The framework diagram goes on to suggest that the risk of heterosexual acquisition of HIV of young people, might be influenced by both household wealth and the individual socioeconomic roles played by young people. These factors might interact with the behavioural pathways articulated in the types of theories described in the previous chapter.

The conceptual framework goes on to articulate explicitly four dimensions of sexual behaviour that might influence the risk of heterosexually acquiring HIV infection, and which might themselves be influenced by the preceding socioeconomic forces. These are whether or not the individual is *sexually active*, the *rate of partner change, the characteristics and risk of infection of sexual partners* and the *use of condoms within partnerships.* Additionally, there are other factors that influence risk of HIV infection. These are grouped as *heterosexual cofactors* and *other routes of transmission.* Heterosexual cofactors include the presence of male circumcision or other STIs. Alternative routes of transmission include iatrogenically acquired infection such as blood transfusion or infection with needles. The framework suggests that these might also be influenced by the level of household wealth and individual socioeconomic roles, although they are not directly under study here. At the distal causal end of the framework lies the risk of HIV infection of young people, which, through a variety of direct and indirect pathways, may be influenced by socioeconomic factors. Implicit in the framework is that different socioeconomic factors might differentially impact different characteristics of HIV risk.

Finally, the framework identifies that there may be other factors that may influence different components of the system. Of particular importance in epidemiological studies are factors that may confound the relationships being studied.

Arrows within the framework diagram articulate the primary direction of enquiry of the thesis. The direction of arrows does not suggest that causality can only run in the direction of the arrows; it is recognised, for example, that aspects of sexual behaviour or the acquisition of HIV infection may influence socioeconomic status variables. Of particular importance in the current thesis will be consideration of the role of childbearing among young women. Pregnancy leading to childbearing is directly the result of unprotected sexual intercourse, and will be considered as an outcome variable in analyses. However, pregnancy and childbearing are likely also to be strong predictors of whether young women remain in school or migrate away from the rural home.

2.04 Summary of chapter

In South Africa, there are high levels of poverty in rural areas, fuelled by high unemployment, entrenched labour migration and poor agricultural potential. Young people approaching adulthood in these areas have high levels of attendance at primary school. However, many do not complete secondary schooling despite high levels of attendance. Upon leaving school, prospects for local employment are very poor, while prospects for employment that requires temporary migration away from the rural home are slightly better. Many young people who leave school remain unemployed, staying either at the rural home or migrating from the home in search of work. Some successfully acquire work and these individuals usually generate a significantly higher income.

During the same age period that young people are negotiating these economic realities of rural South Africa, they are often forming early sexual relationships, primarily occurring outside of marriage. This entry into sexual activity may put individuals into contact with HIV. Practices that increase the spread of HIV, including high numbers of partners for males, high age differentials between partners, and lack of condom use are common in South Africa. HIV prevalence and incidence remain high in the country, with the highest levels of new infection among the young, particularly women.

There are a number of pathways through which the economic background and roles played by young people may affect their sexual behaviour patterns and / or risk of HIV infection. Some influences may increase risk of infection, while others may decrease the risk of infection. The probability of an individual becoming infected with HIV through heterosexual intercourse is dependent on whether they have sex, rates of partner change, and the characteristics of their partners and other factors that impact the risk of transmission from one partner to the other during sexual intercourse. These characteristics of sexual behaviour may be influenced by individual psychosocial, social network and power structures, which themselves may be influenced by socioeconomic forces and the socioeconomic roles played by young people.

Socioeconomic forces might then act to influence many different determinants of HIV risk and might act differently within different population groups. These forces might act on individuals, households or societies. The complex interplay of the presence or absence of socioeconomic influence on characteristics within each of these dimensions will determine the ultimate magnitude and direction of any overall measure of association within adult populations where heterosexual sex remains the main route through which HIV is acquired. The magnitude and direction of this relationship might also alter over time. The following chapter will review the empirical literature on the relationship between socioeconomic factors and risk of HIV infection from studies conducted in southern, eastern and central Africa in order to identify trends and gaps in this literature to inform the generation of hypotheses for the current study.

Chapter 3. Literature review

This chapter describes the findings of a review of the literature on the relationship between markers of household wealth, education, migration or work and sexual behaviour or HIV infection among populations in southern, central and eastern Africa. The findings of the review are summarised in relation to the conceptual framework described in the previous chapter. Trends and gaps in the literature are identified and hypotheses for the current study developed.

3.01 Search Strategy and inclusion criteria

The major medical research databases (Embase and PubMed) were searched at the beginning of the project, and the review was last updated in late 2005. Boolean syntax were used to combine MEsH terms for "sexual behaviour" OR "HIV prevalence" OR "HIV infections/epidemiology" with (AND) a term for "Africa, south of the Sahara", AND terms including "educational status" OR "transients and migrants" OR "unemployment" OR "employment" OR "social class" OR "occupations" OR "socioeconomic factors". Data from western and northern Africa were excluded since the characteristics of the HIV epidemic were felt to be significantly different in these regions. Approximately 600 articles were identified in this manner. The titles and abstracts of articles identified were reviewed. In addition, reference lists of all relevant articles were searched, including those of a number of systematic review articles. Some unpublished reports of large studies where some component of the study had been published or peer reviewed were also included in the review.

Articles that contained data on the association between socioeconomic factors (such as household wealth, educational attainment or current schooling, circular migration, travel or mobility, or individual income or employment) and either sexual behaviour or HIV-1 infection were reviewed. Table 2.3 contains details of articles identified in this manner that contain quantitative data on these associations. Articles included in this table are those with data on more than 500 individuals (of either sex, or pooled if pooled analyses were presented), conducted in populations broadly representative of the general population, and from countries in southern, eastern or central Africa. Articles containing quantitative data on army recruits who underwent a physical examination, family planning clinic attenders, hotel and bar workers, commercial sex workers, soldiers, truck drivers, STI clinic patients, monogamous married men and hospitalized patients were excluded since these groups were not likely to be representative of the general population. Studies were only included if they appropriately adjusted for potential confounders. This meant that all results had to be controlled or stratified for the effect of gender. Additionally, analyses in which HIV was the outcome also had to control for at least for age and setting (urban or rural) in the presentation of their results since these were felt to be significant potential confounders. In one included article a dummy variable for age group was entered to, but not included in, a stepwise regression model (146). In a previously published review on the link between educational attainment and HIV, articles were also excluded from review if they "over-adjusted" multivariate analyses for indicators that might be expected to mediate the relationship between socioeconomic factors and HIV infection (particularly sexual behaviour or STI symptoms / serology) (147). However, in many cases where this occurred it was not possible to isolate the impact of adjustment for these factors from the impact of adjustment for other true confounders. Other recent analyses have explicitly adjusted for these pathway variables and in such cases this "overadjustment" had limited impact on the measure of effect presented (147-149). Articles presenting overadjusted analyses were not excluded from the review presented here, although these analyses are identified in the table and the text. Finally, a small number of articles that contained interesting data, but that did not meet the inclusion criteria, have been commented on in the text, and the reason for their non-inclusion clearly identified in each case.

3.02 Results of the review

(a) Overview of studies included in the review

Quantitative data were available from general population studies, antenatal clinic surveillance systems, cohorts of workers and students. Where general population surveys had been conducted most had adopted household cluster sampling techniques which were well described. Less well described in most papers were the criteria used to determine household membership or residence. However, some studies clearly adopted a '*de jure*' criteria of household membership that did not require current residence at the time of enumeration (143, 150, 151), while others adopted a '*de facto*' approach requiring individuals to be currently sleeping in the home (149, 152, 153). Most studies were conducted within clearly identified regions though a small number were nationally representative (see Table).

Table 3.1 : Studies reporting data on the association between household wealth, education, migration or employment and sexual behaviour or HIV among populations representative of the general population in sub-Saharan Africa published 1990-2004.

Reference(s)	Study Popn	Study Design	Country	National / regional	Setting	Year of data collection	No.&Sex of Participants			Age	Socioeconomic factors				Sexual behaviour markers	Data on HIV
							M	F	P		HH Wealth	Educn	Mobility	Occupn/Income		
Allen (146)	Pre/post	CS	Rwanda	R	U	85-86		b		19-37	Ed	AT	Mov	Inc		(X)
Glynn (154)	Pop	CS	Malawi	R	R	87-89			c	15+	Hou	AT		JT		X
Crampin (155)	Pop, ANC	SCS	Malawi	R	R	88-00,91-93,98-01			a/b	15-54		AT	Res	JT		X
Dallabeta (156)	ANC	CS	Malawi	R	U	89-90		c		13-45	Fac, Ed	AT		Emp		(X)
Chao (157)	ANC	CS	Rwanda	R	R	89-91		c		<19->30	Inc	AT				(X)
Kwesigabo (158)	Pop	SCS	Tanzania	R	U,R	87, 93, 96			b	15-54		AT		JT		X
GPA (159, 160)	Pop	CS	Multiple	N	M	89-91	b/c	b/c		15-49		AT		JT	NRP, Con, Com	
De Walque (161, 162)	Pop	CS, Coh	Uganda	R	R	90to00			b	13+		AT			Con, Life, Rec	X
Barongo (148)	Pop	CS	Tanzania	R	M	90-91	b	b		15-54		AT	Tra, Mov	Emp, JT		X
Wannan (163)	ANC	CS	Zaire	R	M	90-91		b			JTH	AT				X
Senkoro (164)	Work	Coh	Tanzania	R	U	91to96	b	a		<25->45		AT				X
Serwadda (151)	Pop	CS	Uganda	R	R	89	a	a		15-59		AT	Tra	JT		X
Rakail (165, 166)	Pop	CS	Uganda	R	R	90	b	b		13+		AT				X
Djamba (167)	Pop	CS	Zaire	R	U	91		a		Mean 33yrs		AT			Pre	
Mwanza (168, 169)	Pop	CS	Tanzania	R	R	91-92	c	c		15-54		AT	Tra	JT		X
Kilian (170)	ANC	SCS	Uganda	R	U	91-94, 95-97		b		15-49		AT				X
Rakaill (171)	Pop	CS	Uganda	R	R	92			b	13+	Ass	AT		JT		X
Mnyika (172, 173)	Pop	CS	Tanzania	R	M	92		a	b	15-54		AT	Tra	JT		X

Reference(s)	Study Popn	Study Design	Country	National / regional	Setting	Year of data collection	No.&Sex of Participants			Age	Socioeconomic factors				Sexual behaviour markers	Data on HIV
							M	F	P		HH Wealth	Educn	Mobility	Occupn/Income		
Lugoe (174)	Students	CS	Tanzania	R	M	93			a	13-25	Sch	AT			Ever, Con	
Bassett (175)	Work	CS	Zimbabwe	R	U	93-95	b			18+	Hou	AT		JT		(X)
DHS (137, 176)	Pop	CS	Multiple	N	M	93/98	b	b		15-49	Ass	AT, In		JT	Ever, NRP, Con	
Fylkesnes I (177)	ANC	CS	Zambia	N	U,R	94		c		15-44		AT				X
Morris (143)	Pop	CS	Uganda	R	R	94			b	15-49		AT	Tra	JT	Con, NRP, AD, Com	
Rakai Trial (178)	Pop	Coh	Uganda	R	R	94to99		C		15-49		AT				X
Bloom(179)	Pop	CS	Tanzania	R	R	94-95	b	b		18-59	Ass	AT	Res	JT		X
Fylkesnes II (180)	ANC	SCS	Zambia	N	U,R	94-98		c		15-39		AT				X
	Pop	SCS	Zambia	N	U,R	96, 99	b	b		15-49		AT, In			Con	
Meekers (181)	Pop	CS	Botswana	R	U	95	b	b		13-18		AT, In			Ever	
Fontanet (182)	Work	CS	Ethiopia	R	U	95-96			b			AT	LM			
Kapiga (183)	Pop	CS	Tanzania	N	M	96			c	15-49		AT			High	
Four cities study (149, 152, 153)	Pop	CS	Kenya	R	U	97-98	a	a		15-49	Mix	AT, In	Tra	JT	Con, Life, Rec, Num, AD, Ever, Dry	X
	Pop	CS	Zambia	R	U	97-98	a	a		15-49		AT, In	Tra	JT	Con, Life, Rec, Num, AD, Ever	X
Gregson2(150)	Pop	CS	Zimbabwe	R	R	98			a	17-54M, 15-44F	Liv	AT	Res	Emp		X
Camlin (56)	Pop	CS	South Africa	N	M	98		c		15-49		AT			Con	
Zuma (184)	Pop	CS	South Africa	R	U	98		b		13-60			LM			X
Magnani (185)	Pop	CS	Zambia	R	U	98			b	10-24		AT		Inc	Ever, Life, Rec, Con	

Reference(s)	Study Popn	Study Design	Country	National / regional	Setting	Year of data collection	No.&Sex of Participants			Age	Socioeconomic factors				Sexual behaviour markers	Data on HIV
							M	F	P		HH Wealth	Educn	Mobility	Occupn/Income		
Lurie (144, 186)	M/NM	CS	South Africa	R	R,U	98-00							LM		Rec, Con	X
Koenig (187)	Pop	CS	Uganda	R	R	98-99		b		15+		AT			Coe	
Gregson3 (135, 188)	Pop	CS	Zimbabwe	R	R	98-00	b	b		15-39		AT, In	LM, Mov		Con, Life, Rec, AD	X
Mbulateiye (189)	Pop	CS	Uganda	R	R	99-00			c	13+		AT				(X)
Transitions (190, 191)	Pop	CS	South Africa	R	M	01	b	b		14-24	Ass, Ed	AT, In		Emp	Ever, AD, Coe, Com, Con, 2Ab, Rec	
HSRC (45)	Pop	CS	South Africa	N	M	01			c	15-49	Ad					(X)
Kimuna(192)	Pop	CS	Zambia	N	M	01-02	B			15-49		AT		JT	Extra	

Key

Study Population ; Pre/Post = Women from pre and post natal clinics, ANC = Antenatal clinic attenders, Work = Workers, Pop = General population sample, M/NM=Migrants and Non-Migrants.

Study Type; CS=Cross-sectional, SCS=Serial Cross-sectional, Coh= Cohort.

Regional / national; R=Regional, N=Nationally representative

Setting; U=Urban, R=Rural, M=Mixed population, U,R=Urban and rural populations reported separately.

Year; xx-xx denotes years of a single data collection period; xx, xx denotes separate years of data collection, xxtoxx denotes follow up period for a cohort study, xx/xx denotes multiple surveys conducted within the given period

Number and sex; F=Females, M=Males, P=Pooled sexes. a=500-1000 participants, b=1000-5000 participants, c= >5000 participants

Household Wealth; Hou=Housing quality, JTH = Job type of sexual partner or husband, Ass = Asset score, Fac = Access to facilities, Ed = Education of husband / household head, Inc = Income, Mix = Combined indicators, Liv=Livelihood class, Ad=Self report of adequacy of household income.

Education; AT=Educational attainment, In=Attendance at School, Sch = School type

Migrancy / mobility; Tra=Travel, Res=Marker of recent residence in study area, LM=Labour migrant, Mov=Recently moved.

Occupation / Income; JT=Job Type, Emp=Employment or Unemployment, Inc=personal income

Sexual behaviour markers; Ever=Ever had sex or age of debut, Con = Marker of condom use, NRP = Indicator of having sex with a non-regular or non-spousal partner, Life = Lifetime number of partners, Rec=Recent number of partners, AD=Age difference with partners, Dry=Dry sex, 2Ab = Secondary abstinence, Coe=Coercive sex, Com=Commercial sex or sex with resource exchange. High = Composite marker of high risk sex, Extra=Extra marital sex

HIV; X=Data on HIV included in paper, (X) Data on HIV in overadjusted analyses included in paper.

The earliest available data was collected in 1985-6, while the latest studies reported in the literature contained data collected during 2001. More data was available on females than on either mixed or male only populations. Three studies reported data on incidence of HIV infection from a cohort study (161, 164, 178), two reported serial cross-sectional surveys (155, 158), while a further two explicitly investigated trends over time from antenatal clinic data (170, 180). Studies were approximately evenly distributed between rural and urban populations, with a small number providing data on both or mixed populations. Data came from only a selection of countries, with data from more than one study available for Malawi, Rwanda, Tanzania, Uganda, South Africa, Zimbabwe and Zambia. No data was available for a number of countries in the region. In describing the results of the review in the sections below, studies presenting data on South Africa are discussed in detail.

Educational attainment, measured by grade or level attainment or number of years attending school, was the marker of socioeconomic status most commonly reported (see Table 2.1). Only four studies examined the association between current attendance at school and HIV risk outcomes. Household wealth measures were reported by a number of studies, and this was measured using a variety of rapid indicator approaches, focusing on housing quality, ownership of assets, household or household head income, self reported adequacy of household income or socioeconomic stratum. In one study a mixture of individual and household level markers were combined (147). Markers of employment most often used were job type or unemployment, while personal income was measured in two studies. Migration or mobility was measured in terms of numbers of trips or length of time away from the home, currently sleeping at the home, time living in an area, inability to be available for interview at a previous time point, and one study recruited rural labour migrants at their remote location from the rural home. Some studies reported on a number of different socioeconomic factors. However, none of the studies explicitly presented analyses on the relationship between these factors. Data were either presented separately for their relationship with outcome variables, simultaneously entered into a multivariate model, and / or stepwise regression was used resulting in the inclusion of some but not other factors. Most papers reviewed contained a brief description of the broad socioeconomic context of the study setting, at either local or national level. More detailed descriptions of socioeconomic context were rarely held in the papers reviewed for this study. For most studies, data were presented on socioeconomic risk factors alongside other risk factors, and the presentation of these data was not the primary aim of the paper. Perhaps as a result of this, only a few authors provided a description of the theoretical perspectives underlying their investigation of these associations.

HIV infection was a more commonly reported outcome than measures of sexual behaviour. Where sexual behaviour was reported as an outcome, this was most commonly in the form of condom use with a recent partner, sex during the last 12 months with a non-spousal partner, or lifetime number of partners. A number of studies reported analyses on the relationship between socioeconomic factors and both sexual behaviour and HIV as outcomes, although only two explicitly explored how these relationships were linked (147, 149).

Most analyses reported on adult populations from teenage years to 40s or 50s, adjusting analyses for age. A subset of studies explicitly examined associations among younger age groups. The rationale for this was both because of a specific interest in young people as a high risk population, but also because among young people, where infection, if present, is more likely to have been recently acquired, it is possible to make greater inference about the influence of socioeconomic factors on risk of new HIV infection in the absence of true incidence data. In cases where young people were not the specific focus, a high proportion of respondents had generally been married. In discussing the results of the review below, greater detail has been provided on studies where data focusing on young people (those under 25 years) was presented.

(b) Studies with HIV as an outcome

1) Education

Studies of the association between educational attainment and prevalent HIV infection among groups representative of the male general population, and reporting analyses stratified or adjusted for at least age and setting, tended to report that increased educational attainment was linked to increased prevalence of HIV infection (165, 168, 179) or no association (149, 151, 169, 175, 186). Among females, the pattern was similar, with some studies reporting a positive association (157, 163, 165, 168-170, 177, 179, 180) and some reporting no significant association (146, 149, 156, 172, 186). Finally, in pooled sex analyses adjusted additionally for sex, there was also usually a positive association (154, 171, 173), or no trend (150, 155, 189). There were few exceptions to this pattern. One study reported the opposite association among a group of sugar estate residents of both sexes in Ethiopia (182), while in a study in rural Zimbabwe, there was no association among males, but females with secondary education were less likely to be infected than those with primary only (135). In Masaka, Uganda, there was no association between educational attainment and prevalent HIV infection in 1989-90 in pooled sex analyses, but a significant negative relationship among the young in 1999-2000 (161).

No data from prior to 2000 was available on this issue for South Africa. The national HSRC study conducted in 2001 found higher levels of HIV in higher educational groups in a pooled sex, unadjusted analysis (45). However, this effect was apparently attenuated by adjustment for other sociodemographic factors including age, setting and province. Another study, which focused on migrants from KZN Province found no association between educational attainment and HIV infection rates among males (186).

Where data was available to compare directly, there was some evidence to suggest that the positive association between education and HIV risk was stronger among rural than urban populations (158, 165, 177). The magnitude of the relationship was similar in males and females (165), or stronger in females (169). In studies in Rwanda and Malawi, HIV risk was not associated with a woman's education, but was associated with higher educational status of her partner (146, 156).

A number of studies, or series of analyses, presented data that allowed the comparison of the relationship at different time points. Where this was possible, these studies suggested that a positive relationship between education and HIV was weakening over time (155, 158, 161, 170, 180). In Zambia and Uganda, it was suggested that this was primarily due to greater falls in HIV prevalence among the most educated groups (170, 180). However, among Zambian antenatal clinic attenders (180) and a rural Malawian population of both sexes (155), there was a rise in prevalence among the least educated group over time. One study presented unadjusted analyses that suggested no change in the strength of the relationship over time in a female population in Malawi (193) (study not shown in the table).

A number of studies presented data on young people. These studies reported a positive association between educational attainment and HIV infection among both sexes in Rakai, Uganda (165), no association in either sex in Kenya and Zambia (149), and one reported some evidence of a negative relationship among female Zambian youth (177). A small number of studies reported on the association between current school attendance and HIV infection. In a national Zambian study and a rural Zimbabwean population, HIV prevalence was higher among those out of school (135), while in Kisumu, Kenya and Ndola, Zambia in 1996, no association was noted between attendance at school and HIV infection (149). Details of each of these analyses was provided briefly in the text of the paper, and at least two referred to the potential for bias in this relationship by the impact of childbirth on attendance at school, although they did not attempt to adjust for either childbirth or current pregnancy in analyses.

Three studies reported data on incidence of HIV infection. In a Tanzanian factory cohort, there was no association between educational attainment and risk of incident HIV infection among males, but among females there was a significantly higher level of new infection among women with higher levels of education (164). In Masaka, Uganda, the incidence of infection was lower among more educated females later in the cohort. Among males and older females there was no significant relationship between education and incidence of HIV at any time (161). In Rakai, Uganda, education was not associated with incident infection among women (178).

2) Household wealth

Some studies reported adjusted analyses of the association between indicators of household wealth and HIV infection. Among males, studies reported no association (147, 165) or a positive association (179), while among females, studies reported either no association (156, 165) or a positive relationship (157, 179). In pooled sex analyses there was no significant association noted in any study (45, 154, 155, 171). One study reported that home ownership, as opposed to renting accommodation, was associated with a reduction in risk of HIV infection among male factory workers in Zimbabwe (175). In a subgroup analysis of a Kenyan population in 1996, a composite marker of wealth was negatively associated with risk of HIV infection among young women (15-25 years) but not among older women (25-49 years) or among men of either age group (147). One South African study did present data relevant to this question. The HSRC study found that those whose income allowed for purchase of "most things" or "luxuries" had lower prevalence of HIV than among those whose income was "adequate" or "not enough", with borderline significance in presented analyses (aOR, for most compared to not enough 0.5, 95% CI 0.3-1.0, for "luxuries" compared to "not enough" 0.6, 0.3-1.2, after adjustment for age, sex, setting, province, marital status, race and self reported history of an STI) (45). No other studies reported the association between household economic status and HIV in different settings, at different time periods or among young people alone.

3) Employment or personal income

A number of studies reported the association between HIV infection and employment status or personal income. In some studies, individuals were categorised on the basis of employment categories. Employment types associated with a higher level of cash income or a higher level of mobility from rural areas were generally grouped together. Individuals in these groups had higher levels of HIV infection among both sexes (154, 155, 171, 179) or women alone (148). Other studies reported no association in adjusted analyses between personal income, job type or employment and HIV infection among women (146, 151, 172, 186), men (151, 175, 186) or in pooled sex analyses (150, 173). A national South African study presented unadjusted analyses suggesting no association between employment and HIV prevalence (45). In contrast, another paper presenting unadjusted analyses of the relationship between work grade and HIV infection among industry workforces in three southern African countries (South Africa, Botswana and Zambia) during 1998-00 reported a strong relationship such that those in contract, unskilled or semiskilled positions had higher levels of HIV infection than employees in skilled or management positions (194). However, in the absence of adjustment for age, setting or sex (although the workforces were predominantly male) it is difficult to interpret the results presented. One included study presented data on the relationship between employment type and HIV across three studies conducted in the same area of Malawi among a pooled sex group. Being in a non-agricultural job type was positively associated with risk of HIV infection in 1989-1999. This relationship was somewhat weakened, but still statistically significant in 1998-2000 (155). Conversely, in Tanzania, job type was not associated with HIV infection in rural areas in either 1987 or 1996 among a pooled sex group, while in urban areas traders were at a higher risk of infection in 1987, and this relationship was strengthened in 1993. In Mwanza, Tanzania, in 1990/91, an association between job type and prevalent infection was seen for women but not for men (148). In one presented study, occupation of the sexual partner of women attending antenatal services in Zaire was strongly predictive of the woman's HIV status, with those with partners who were soldiers, policemen or drivers being at much increased risk (163).

4) Migration

Some studies presented adjusted analyses on the relationship between mobility or circular migration and HIV infection. In many cases, increased mobility or being a migrant worker was associated with increased risk of HIV infection (147, 154, 155, 168, 184, 186, 188) among male and female populations, or among females but not males (148). The effect was seen in both urban and rural populations. Migration out of the country, but not local migration was associated with HIV among women in a Tanzanian study (172). However, a smaller number of other studies reported no association between mobility and HIV among women (146, 151, 179), men (151, 179) or in pooled sex analyses (182). One study in a rural area of Malawi suggested a decrease in the magnitude of the association between mobility and HIV infection over time (155). Studies with data on migration or mobility often focused on populations with higher numbers of individuals of greater ages.

5) Multiple socioeconomic markers

A small number of studies reported the relationship between HIV and more than one socioeconomic factor. Most commonly, this included education and some combination of the other dimensions. It was difficult formally to assess how these factors were related from presented analyses. Nevertheless, in some cases the direction of association of educational and occupation factors was similar (155, 158, 163, 166, 171). Education appeared more influential than job type in others (160, 175). In other analyses, household wealth or income was more important than other factors (150, 156, 157, 175). Finally, some analyses emphasised the importance of mobility over job type or education (148, 168, 173). Simultaneous adjustment of multiple SES factors tended to lower the associated odds ratios for each, although only in one case was it possible to separate this from the impact of adjustment for other confounders or pathway variables (154). In this case, simultaneous adjustment lowered the odds ratios associated with increased education and occupational type so that these factors were no longer significant, while the odds ratio associated with housing quality, as a proxy for household wealth, increased slightly.

(c) Studies with sexual behaviour characteristics as an outcome

1) Education

In some studies increased educational attainment was associated with later age at sexual debut among females (135, 137, 149, 176), while in others no association was seen (174, 181). Among males, studies suggested later age at sexual debut among the more educated (135, 149), no relationship between age at sexual debut and educational attainment (137, 174, 176) or a greater chance of having been sexually active among males in secondary education (181). In one pooled sex analysis, sexual debut was less common among those who had higher levels of attainment (185). In three studies it was also shown that younger individuals in school were less likely to be sexually active than those out of school. The potential for early sexual activity to precipitate school leaving was recognised in the discussion of these studies, although no further adjustment was made (137, 181, 191). In another study conducted in Tanzania, having been sexually active was not associated among either males or females with type of school attended (public or private) (174). In urban Botswana, males with secondary education were more likely to be sexually active after adjustment for age and other factors, but there was no association with school attendance, while among females a strong association was noted between attendance at school and ever sexual activity, ascribed once again to the impact of pregnancy on schooling (181).

In Zaire in 1991, educational attainment was associated with a greater chance of having pre-marital sexual intercourse among females but not males, among whom pre-marital sexual intercourse was almost universal (167). Other studies have suggested that having sexual intercourse during the previous year with a non-regular partner was positively associated with educational status among females in rural but not urban areas (137, 176). Among males, a positive association in this relationship was seen in the GPA surveys conducted in multiple countries in 1989-91 (159, 160), but not in DHS data from the mid 1990s (137, 176). Education was not associated with extramarital sex among married Zambian men in 2001-2 (192).

Increased educational attainment was associated with a lower number of recent or lifetime partners among pooled sex samples (180, 185), or was not associated with this outcome in either sex (135, 149). In a pooled sex analysis from Masaka, Uganda, individuals with higher levels of education were more likely to have had a higher number of partners during the previous year (161). In other studies, education was associated with a higher number of lifetime partners among women (185), no association was seen in either sex (135), or a lower number of lifetime partners among women in one city (Kisumu) but not another (Ndola) (149). Among males, there was no association between educational attainment and lifetime numbers of partners in the same study, and there was additionally no association between having had sex on the same day as meeting a partner for either sex in two different urban settings (149). In Tanzania in 1996, females with incomplete primary education were the most likely to have high risk sexual behaviour with this being lowest among those with secondary education and above. In this study high risk behaviour was defined as having other partners if married or having multiple or non-regular partners if unmarried (183).

Higher educational status was associated with a greater likelihood of condom use in most studies where this was reported among both males and females (56, 137, 149, 152, 159, 171, 174, 176, 180, 185). In one population in Kenya, the female partner's educational level was positively associated with condom use (152). In a smaller number of studies no association was noted between education and condom use in either sex (135, 143, 160), but no studies reported greater levels of condom use among less educated groups.

In rural Uganda, women with primary education were more likely than those with either secondary or no education to have experienced coercive sex (187). Lower educational attainment was associated with slightly larger age gaps with partners among females in a rural Zimbabwean setting (135), but was not associated with age differentials between sexual partners for partnerships reported by females in two urban settings (149). For males, more educated individuals were more likely to report partnerships characterised by a wide age gap in one city (Kisumu) but not another (Ndola) (149). Data from the GPA surveys in 1989-1991 suggested that more educated males were more likely to have engaged in commercial sex during the previous year (160), although this finding was not seen in a more recent study (149). In turn, DHS data from the mid 1990s suggested that in some, but not all countries, more educated women were more likely to have received resources in exchange for sex, although this was not seen in urban Kenya and Uganda where the opposite relationship was seen (137). More educated women were also less likely to report exchanging money for sex in Kisumu, Kenya and Ndola, Zambia in 1996 (149).

2) Household wealth

Two studies examined the association between household wealth and sexual behaviour in detail. One of these studies was conducted in South Africa while both presented information on young people. The 'Transitions to Adulthood' study conducted in KwaZulu Natal, South Africa (KZN) in 2000 provided useful information on 14-22 year olds. The study was a cross-sectional design, and covered one urban and one rural district. A marker of household wealth was developed through a simple arithmetic measure of asset ownership. The other study was a sub-analysis of data collected in Kisumu, Kenya as part of the "Multisite study on the heterogeneity of HIV epidemics in African cities". This study was also cross-sectional, and a marker of household socioeconomic status was developed by combining information on access to services, and education and employment of the household head.

Among young males in KwaZulu Natal, greater wealth was not associated in multivariate analyses with the probability of ever having had sex and was not associated with secondary abstinence (190). Among 14-22 year old females in the same study, greater wealth was associated with a lower probability of ever having had sex. Greater wealth was associated with a lower chance of having had more than one partner in the last 12 months among women but not men (190). In the Kenyan population, older, wealthier males reported more lifetime partners (147). Among young males (14-25 years) there was no association between a marker of wealth and any sexual behavioural characteristics, while among the older group, wealthier males also reported later age at first intercourse and marriage. Among females in the Kenyan study, greater wealth was associated with later age at first intercourse and marriage, but not with lifetime number of partners (147).

Among males in the South African study, being in the highest wealth quintile was also associated with the probability of the age gap with the most recent partner being smaller compared to the poorest group (190). Among females in the same study, greater wealth was not associated with age differential in partnerships, but greater wealth was negatively associated with ever having experienced forced sex or having traded sex for money (190). In the Kenyan study, greater wealth was associated with smaller age differentials with marital partners in the group of females aged 25-49 years, but not among younger females or males (147).

Greater wealth was not associated with condom use at last sex among young males, but among 14-22 year old females in South Africa, greater wealth was associated with a greater likelihood of condom use at last sex (190). In the Kenyan urban population, wealthier males aged 25-49 had greater levels of condom use and lower levels of HSV-2 infection, though this was not seen among 15-24 year olds (147). Greater wealth among females was also associated with lower levels of 'dry sex' and HSV-2 infection among both 15-24 year olds and 25-49 year olds, and higher levels of condom use. Condom use was also associated with greater household wealth in a number of other adjusted (152), and unadjusted analyses (165, 171), including an unadjusted study conducted among South African women (195).

3) Employment or personal income

A small number of studies examined the association between individual income or occupation and sexual behaviour characteristics. Higher personal income was associated with a greater number of sexual partners, but not with condom use, in a pooled-sex study in Zambia (185), while extramarital sex among married men was not associated with job type (192). Individuals in professional employment and with higher income had greater levels of condom use among men in Ndola, Zambia, but no relationship was seen among either sex in Kisumu (152). Among a pooled sex group in Rakai, Uganda, professional employees were more likely to report condom use with non-spousal partners than farmers, labourers or students (143). In DHS surveys conducted during the mid 1990s in a variety of settings, women in manual employment were more likely to have had sex with a non-spousal partner during the previous year than those who were either in skilled employment or without a job, while among males the relationship was less clear (176). In a South African study men who were currently working were the most likely to have been sexually active in the past year, but were also more likely to have used a condom, while no association was seen among young women (191).

4) Migration

Finally, a small number of studies examined the association between mobility and sexual behaviour. In South Africa, data has been presented to suggest that migrants were more likely to have a higher number of recent partners and not to use condoms, or to use condoms less frequently with regular partners than non-migrants (184, 196). Detailed epidemic modelling work using data from South Africa, has, however, challenged the popular notion that returning male migrants are those most likely to infect their partners at home, showing that both migrants and their current partners were more likely to have been infected by someone other than their current main partner (144). In Rakai, Uganda in the mid 1990s a study suggested that those travelling away from this rural area had more positive attitudes to condoms, used them more often, and did not have higher levels of partner change or a riskier partner profile (143). This study also suggested that the partners of migrants of either sex were likely to be younger and more educated, and male migrants reported partnerships of shorter duration. Travellers did not apparently report partnerships that were more assortative with regard to higher risk characteristics or with sex workers than non-migrants. In another study, travel was not associated with frequent condom use among females in two cities, while among males there was no association in Kisumu, but in Ndola those who had made more than three trips in the past 12 months were significantly more likely to report frequent condom use (152). Migrant status has also been associated with lower chance of male condom use with regular partners only (not marital or casual partners) in one unadjusted analysis (186). No studies identified in this review examined mobility or migration and its association with sexual behaviour among young people alone.

3.03 Trends and gaps in the empirical literature

Empirical data suggested that risk of HIV infection was associated with traits including high levels of mobility, education and wealth in southern, central and eastern African countries during the late 1980s and early 1990s. However, later data suggested a shift, at least in some countries, towards this relationship disappearing or perhaps reversing. There was evidence from a number of studies that some protective sexual behaviours, in particular the use of condoms, were associated with higher levels of education and wealth. However, the relationship between economic factors and sexual behaviours was complex, with other behavioural risk factors perhaps being more common among higher SES individuals.

Three major gaps in the literature were identified in the review to which it was felt the current study could contribute new information. Firstly, the social epidemiology of sexual behaviour and HIV infection in South Africa remained unclear. Secondly, the association among young people had been under-studied. Thirdly, few studies had explicitly tried to investigate the mechanism of the association between socioeconomic factors and HIV infection. Each of these limitations in the literature is discussed in more detail below. The chapter ends with a statement of the specific hypotheses to be tested in the current research.

3.03.1 The social epidemiology of sexual behaviour and HIV infection in South Africa

The association between socioeconomic factors and sexual behaviour or risk of HIV infection in South Africa remains unclear. Only a small number of studies have reported data from South Africa. A larger number of previous studies have been conducted in sub-Saharan Africa that have examined the association between socioeconomic factors and HIV risk characteristics. However, information from these studies cannot safely be generalised to young people in South Africa for at least two reasons.

Firstly, the epidemic of HIV in South Africa has been different to that seen in other countries. The South African epidemic has been particularly severe, but rises in prevalence were not seen until much later than in many countries in eastern and southern Africa. Because of this later epidemic, the environment in which HIV has spread in South Africa has been different to other settings in the region. South Africa's response has been characterized by poor implementation of policy and perceived lack of political will to deal with the issue (197, 198). In this respect, South Africa's response has not been dissimilar to that of other countries in the region. However, it is likely that overall levels of HIV awareness were nevertheless higher in South Africa during the 1990s than in eastern and central African countries during the early part of the 1980s when HIV prevalence was rising there, since global awareness of the problem was much stronger during this later period. Additionally, South Africa had a national AIDS plan by 1994 when prevalence was still below 10%, has characteristically had a strong NGO sector that was active from early in the epidemic and has a relatively well developed communication infrastructure. Even if the response to the HIV epidemic has been inadequate, awareness of the existence of HIV has been high during much of the epidemic phase of spread. This might have had consequences for the social dynamics of epidemic spread of HIV in this country (129). Over and Piot (1996) suggested that the correlation of high socioeconomic status with high prevalence of HIV seen in studies conducted in the 1980s in eastern Africa might fade over time in any given setting (129). Indeed, there is some evidence that decreases in HIV prevalence in Uganda and Zambia over time have been associated with a tendency to move away from higher levels of prevalent infection in higher SES groups (161, 170, 180). It remains uncertain to what extent this represents a change in incidence patterns or sexual behaviour, since this may also be affected by patterns of mortality from HIV/AIDS. However, behavioural theories also suggest that in the presence of HIV prevention campaigns the educated and wealthy will be most likely to change their behaviour (199). It is possible then that the pattern seen in countries where HIV prevalence rose most rapidly during the 1980s may not be seen in South Africa. In Thailand, where some success has been had in controlling the population level spread of HIV, the epidemic also emerged in an environment of greater global and local awareness. In Thailand, a negative gradient of HIV infection across educational attainment categories has been seen since the early 1990s (200).

Secondly, South Africa is a wealthier country than most other countries in the region, but with a history that has led to a unique socioeconomic environment. South Africa is a middle income country characterized by massive income inequality (8). Much of the economy is cash based and young people are surrounded by images of wealth and modern living. Many young people in South Africa are acutely aware of the resources that poverty denies them. The "rural" former homeland villages in which this study was conducted might have considerable differences to other areas described as conducted in rural areas. In South Africa, such areas have high population density, and while many households engage in agricultural activities, few are able or desire to support themselves on this basis. Migrant labour forms a particularly important component of the rural economy in South Africa, although unemployment is very high, particularly among the young. In contrast to many other rural areas in sub-Saharan Africa, travel is highly accessible and probably not limited to wealthier members of society. Unlike many sub-Saharan African settings, South Africa has almost universal completion of primary school and widespread secondary school enrolment, with high levels of literacy (201, 202).

These issues relating to South Africa make it difficult to generalise the results of previous research. What should be noted, however, is that given prevailing theory it might be expected that both of these factors might make it less likely that higher socioeconomic position would be associated with a greater risk of HIV infection in South Africa. This is because it has been suggested that HIV prevention messages were probably more widespread in South Africa during the early spread of HIV, and that this may be likely to cause wealthier, more educated individuals to change their behaviour first and because mobility, which has been associated with increased wealth and high risk of HIV infection in many rural settings, may not be so related to wealth in rural South Africa.

3.03.2 The social epidemiology of sexual behaviour and HIV infection among young people

Only a few studies have previously examined the association between socioeconomic factors and risk of HIV infection specifically among young people, many of whom remain unmarried. There are a number of reasons why this group may be different from adult populations.

Firstly, previous studies have often used data from antenatal clinic surveillance systems. However, these data may be particularly biased in estimating HIV prevalence among young people, many of whom remain sexually inactive.

Secondly, where HIV prevalence is the outcome investigated, results from population studies among younger groups may more closely mirror the association with incident infection. Younger populations are likely to have started sexual activity more recently, and so HIV infection is more likely to have been acquired recently, and consequently less likely to have caused death. This is particularly important since the relationship between social status and risk of HIV infection may change over time. It is also important because HIV-associated mortality will be likely to cause a decrease in HIV prevalence among older groups who have been infected for a longer period. Mortality rates may differ between socioeconomic groups since higher socioeconomic individuals may have better nutrition, a lower risk of contraction of AIDS defining illnesses such as TB or pneumonia, and better access to treatment. However, if indeed higher socioeconomic status groups were infected earlier in the epidemic, AIDS-associated mortality may be higher among this group in the absence of widely available AIDS treatment.

Young people may also differ from adults for reasons other than epidemiological artefact. Adolescence and young adulthood is a period of rapid hormonal change, development of cognitive processes and changing social status. Complex abstract thinking develops only towards the end of the adolescent period (5). In early and middle adolescence, young people may be capable only of relatively concrete thinking and continue to see themselves as "bullet proof". Adolescents and young adults are thus often the focus of targeted prevention campaigns, which should take account of these realities (3), although many do not. Many young people also remain unmarried, and norms around sexual behaviour change in important ways for both partners after marriage.

For these reasons, epidemiological studies need to consider the specific situation of young people. Relatively few previous studies have identified measures of social status made explicitly relevant to young people. For example, previous studies among young people have examined the association between educational attainment and HIV risk characteristics (149, 165, 170, 174, 180). However, the interpretation of educational attainment among young people, especially those who remain in school, is not entirely clear. Among adults who have completed their schooling, level of attainment in education in the past can be interpreted as being indicative of the socioeconomic environment during childhood and of a store of human capital leading to better prospects for employment and more optimism about the future. However, among the young this interpretation is more problematic since many young people remain in school and these benefits have not yet accrued. Similarly, household wealth may be a more important predictor of welfare or perceived social status than personal income in this group, though neither have been often studied. Finally, the association of migrancy with HIV risk characteristics among young people has not been studied to date.

3.03.3 Mechanisms underlying the social epidemiology of HIV infection

The third major gap in the literature identified by this review was that the mechanism of association between socioeconomic factors and sexual behaviour or risk of HIV infection has rarely been explicitly studied using epidemiological data. Investigating the social epidemiology of HIV was the primary aim of only a small proportion of the studies identified in the review. Where an association between socioeconomic factors and risk of HIV infection was noted, and in a few cases where no association was noted, explanations of mechanism were briefly discussed including many of the mechanisms of association highlighted in previous section. However, generally studies did not pre-hypothesise how socioeconomic factors may be linked to HIV risk or perform additional analyses to explore causal pathways, with a few exceptions (27, 147, 149, 153, 162, 171). In many cases this was due to a lack of data since relatively few studies collected data on both sexual behaviour and HIV infection. However, even where studies did collect data on sexual behaviour, with or without HIV data, it was often not made explicit if, or how, certain sexual behaviour traits were hypothesised to be affected by socioeconomic factors, and through what pathways. Finally, relatively few studies simultaneously collected data on multiple aspects of social status, or explored the association between them where they did collect this data. This meant that often associations with educational attainment, which was the most common marker adopted, were interpreted in terms of the impact of this factor on mobility or income, without data being made available on these factors. Due to these limitations, of the multiple pathways through which socioeconomic factors may be associated with risk of HIV infection, it remains unclear which have been validated in research to date.

3.04 Hypotheses

This study will explore the social epidemiology of sexual behaviour and HIV infection among a sample of unmarried young people of both sexes, aged 14-25 years, from a former bantustan in rural South Africa.

Prior knowledge of the setting suggested that there would be a high level of poverty in the study area. It was expected that there would, nevertheless, be clear differences in wealth between households across the study site. It was expected that same-age individuals of either sex from households of higher wealth would be more likely to be attending school or college than those from poorer homes. It was also expected that some young people would be temporary migrants. The primary force driving temporary migration was felt likely to be the socioeconomic environment, including the lack of local labour opportunities, which would affect all young people. However, migrancy was also expected among some young people for the purposes of education.

Individuals attending school were hypothesized to be more likely to have higher levels of HIV awareness (knowledge, communication, testing and tolerance) than those not doing so due to increased exposure to and engagement with HIV information and prevention campaigns. It was hypothesized that migrants might also have higher levels of HIV awareness than non-migrants, since they would be more likely to be exposed to HIV prevention campaigns in cities and other migrancy destinations.

Individuals in school were expected to be more likely to remain sexually inactive and to have lower numbers of partners due to social norms and pressures from teachers, students and peer groups. Further, it was hypothesised that individuals in school or college might have more closed sexual networks, partners of more similar ages and relationships less likely to involve resource exchange. As a consequence of all of these factors individuals of both sexes attending school were hypothesised to report higher levels of condom use and lower levels of prevalent HIV infection than peers no longer attending full time education.

Despite higher levels of HIV awareness, temporary migrants were hypothesized to report higher levels of sexual activity and partner turnover due to lack of social pressure from household and community members, and supported by social norms for migrant males, and driven by a need for the resources that sexual partnerships can bring for young migrant women. Young people who were migrants were also hypothesised to be more likely to conform to adult norms surrounding sexual behaviour including wide age gaps between sexual partners and higher levels of male to female resource exchange. Migrants were also expected to have more casual or short relationships. Due to these factors, it was hypothesized that young people who were temporary migrants would be less likely to report condom use and have higher levels of prevalent HIV infection than individuals who stayed at home.

As described above, attendance at school was expected to be associated with higher wealth, and also with a lower HIV risk profile. As described earlier, the epidemic of HIV is now well established in South Africa and HIV prevention resources are increasingly available through government and NGO efforts. In line with previously outlined theory, it was therefore also hypothesised that young people from higher wealth households would exhibit lower risk characteristics in this setting. In addition, individuals from poorer backgrounds were expected to have lower levels of HIV awareness, and, further to have riskier sexual behaviour profiles due to lower self efficacy to act on prevention messages that emphasize abstaining, being faithful to one partner and using condoms. Individuals of either sex from households of lower wealth were expected to have higher levels of HIV infection. It was expected that some part of the association between household wealth and these outcomes would be mediated by differences in the schooling profile of young people from different wealth backgrounds.

Thus, the primary hypotheses to be tested in this thesis were;

1) Young people from households of lower wealth would have higher levels of risk of HIV infection than those from wealthier households

2) Young people attending school or college would have lower levels of risk of HIV infection than those who were not attending school or college

3) Young people who were migrants would have higher levels of risk of HIV infection than those who stayed at home

As well as testing these hypotheses, as a secondary aim the study sought to explore the mechanism of any associations noted between socioeconomic factors and risk of HIV infection.

This thesis was conceptualized as being primarily exploratory in origin and will describe a large number of statistical associations. Nevertheless, formalizing specific null hypotheses and assessing the statistical power generated by the study to examine relevant associations was an essential component of the study design. Formal power calculations are presented in Chapter 4 relating to each of the primary exposures and three main high HIV risk outcome measures – high numbers of lifetime sexual partners, no condom use at last sexual intercourse and prevalent HIV infection.

Chapter 4. Study Methods

This chapter describes the objectives and setting of the study, and the methods used to meet the study objectives.

4.01 Study Objectives

The objectives of this study were to;

1. Describe the characteristics of the study sample of young people and their households in the Sekhukhuneland region of South Africa, and assess the representativeness of this study sample (Chapter 5).
2. Generate three measures of household wealth by different methods and assess their distribution and association. One method will be selected for use in the remainder of the study. Following this, patterns of attendance at school or college and temporary migration among young people of either sex who are residents of the study households will be described, and the association of these roles with household wealth investigated (Chapter 6).
3. Describe the distribution of HIV awareness and sexual behaviour characteristics and asses the prevalence of HIV infection and factors associated with this among young men and women (Chapter 7).
4. Describe the association between household wealth and HIV awareness, sexual behaviour and risk of HIV infection among young men and women. Go on to describe the association between the individual socioeconomic roles of young people and these outcomes.

Finally, investigate the extent of interaction between wealth and socioeconomic roles and describe the degree to which specific pathway variables mediate differences in the risk profiles of individuals in different socioeconomic groups (Chapter 8)

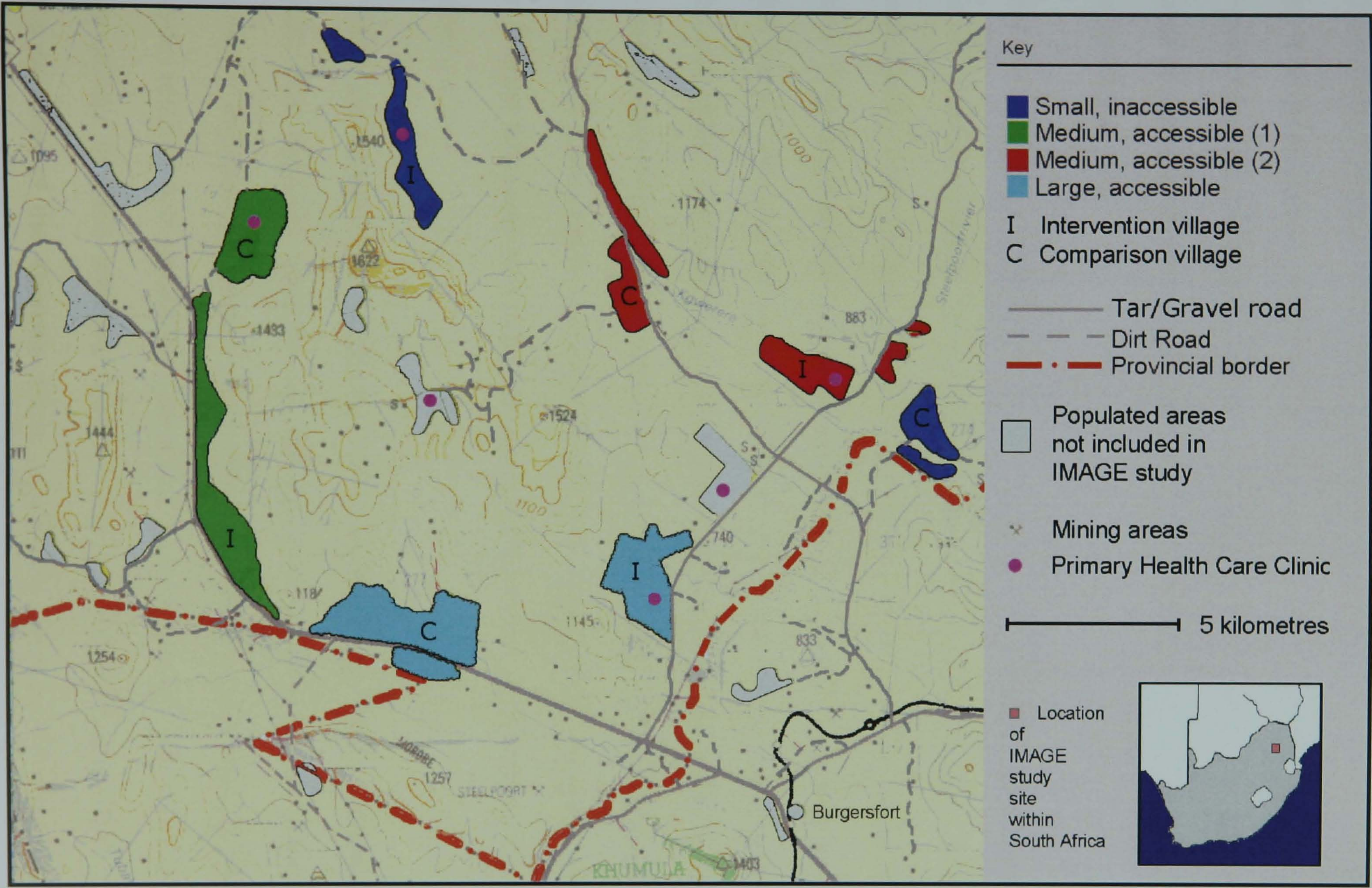
5. Explore the sensitivity of the study findings to assumptions about the validity of the study data (Chapter 9).

6. Discuss the findings of the study in terms of their internal validity and previous research in the field, and discuss the potential implications of the findings for the control of HIV/AIDS in sub-Saharan Africa (Chapter 10).

4.02 Study Setting

The remainder of this thesis describes the conduct and results of an epidemiological study conducted during 2001. The study was set in eight rural villages of Sekhukhuneland, a district of the former homeland region ('bantustan') of Lebowa in South Africa's north east (Figure 4.1). The villages now sit in Limpopo province, just to the north of the border with Mpumalanga province. The region sits on the brow of the Drakensberg escarpment approximately 1000m above sea-level, some 500m above the lowveld to the East, and 600m below Gauteng to the west. A rapidly developing trading centre, which sits just inside the Mpumalanga border, is the major centre of economic activity for residents of all the study villages.

Figure 4.1: Map of the study setting in Sekhukhuneland on the border of Limpopo and Mpumalanga provinces



4.02.1 History

Sekhukhuneland is a region steeped in the history of the BaPedi people. In 1879 the BaPedi kingdom was finally overcome by the combined forces of the British and Swazi armies after many decades of resistance. In the 120 years that have followed, the region's history has reflected broad patterns common to much of rural South Africa while also being the site of a number of unique historical events (6).

As the early decades of the 20th century passed the economic and social order of Sekhukhuneland increasingly rested on migrant labour. While originally many communities continued to plough and herd cattle, the push towards "rural development" encapsulated in the 1936 Native Trust and Land Act left homes with less and less land. Cattle were increasingly culled in the 1930s and 1940s, ostensibly to preserve the land. Labour migrants from the area maintained important links with their rural homesteads. Cultural ties with the bantustan were maintained through regional organisations based in the cities. Migrant labourers provided important remittances to support those back at home.

After 1948, the National Party's election victory meant life for Sekhukhuneland residents and migrants became increasingly strained. The Native Affairs Department increasingly sought to implement *boipušo* (self-rule) in the Reserves, which in turn generated increasing levels of political action among the population. In 1958 tensions erupted in the Sekhukhuneland revolt in which supporters of the imposed Bantu authorities (termed *marangera*, or

Rangers) were attacked and killed by the BaPedi. A small victory was won with the return of the previously deported, but widely respected paramount *Morwamotše*, and his wife *Mankopodi*, to the district capital of *Mohlaletše*.

Yet the 30 years that followed would see the final erosion of rural life in Sekhukhuneland as many had once known it. As the apartheid framework was widely implemented, and residents were increasingly moved out of white farming areas, the population boomed in the reserves. Land and livestock ownership was almost completely eroded. Tribal authorities were rapidly established and the Bantu system took hold. Migrating to work became the main mode of survival for the majority of households, and increasing numbers of women also went in search of domestic work positions in the cities. Children were often left to grow up with their grandparents, and to go through the increasingly accessible but restrictive Bantu education system.

In these increasingly desperate conditions, it is no surprise that as the 1980s progressed, and levels of political and resistance activity increased across South Africa, tensions once again surfaced in Sekhukhuneland. In 1986 there was widespread youth activism, including school boycotts and increasing levels of active resistance. This coincided in the region with a spate of witch burnings in Sekhukhuneland communities, highlighting local tensions and frustrations. Further burnings occurred in 1994 around the time of the first free elections in South Africa, and following Mandela's massive election victory in the region there was much for the new administration to do.

4.02.2 Current context

The eight villages involved in this study now lie within the Greater Tubatse region of the Sekhukhuneland cross boundary district municipality, spread over the Mpumalanga and Limpopo provincial border. Data from both the 1996 and 2001 censuses are available for the municipality region, providing useful population data to illustrate the present day context of the study region.

Between 1996-2001 the resident population of the municipality increased from 230,203 to 270,124, an increase of approximately 17% over the 5 year period. It is likely that much of this increase has been driven by in-migrancy. The town of Burgersfort has been rapidly expanding in recent years, and the area is going to be the site of a number of large new mining projects in coming years.

The major language group in the population remains Sepedi. Grouped data suggest that just over 1/3 of the population fall in the age bracket 15-34 years. Approximately 14% of adults over 20 years had completed secondary school in 2001. Unemployment is rife, with over 60% of the total labour force reporting being unemployed, although it has been suggested that self employment was under-reported in the census. Of those who were employed, the major industries of employment were "community / social / personal professions", "mining / quarrying" and "wholesale / retail". The high levels of unemployment have led to high levels of poverty. In 1996 approximately 35% of the households in the municipality reported no cash income in the census, whilst by 2001 this figure had risen to approximately 43%. The social fabric remains characterised by very high levels of labour migration among both sexes, both to and from the area. Some progress has been made in recent years on other aspects of poverty, yet many of the population remain with limited access to basic facilities. The proportion of households living in shacks or homes made from traditional materials fell from about 40% to 28% between 1996 to 2001, while the proportion of households served by electricity rose from 20% to 47% over the same period. Similarly, the proportion of households having only a pit, bucket or no toilet facility fell from 96% to 81% over this period. However, access to water remains highly constrained for much of the population, with under 20% of households having access to piped water in their home or yard in 2001.

At the time of the study, traditional leaderships still maintained a degree of control in some areas, but transitional local councils were increasingly active in most of the study villages. The area was served by one hospital and one health centre, while four of the study villages had primary health care clinics situated inside their borders. Six of the study villages were widely electrified, while the remaining two small, inaccessible villages did not have widespread access to electricity. State funded primary schools were in all and secondary schools in all but one of the study villages, with children from this village generally walking or getting transport to school in the nearest village. Schools generally charged an annual fee of less than 20USD, with some variation between villages. In addition private schools, charging much higher fees were found near to the local trading and administrative centre. Child migration in order to access better quality schooling was not uncommon throughout the study site.

The prevalence of HIV infection among antenatal clinic attenders in the region of Limpopo province where the eight villages in this study are located was

13.2% in 2000. This figure reflects the situation across the largely rural Limpopo province, in contrast with the more developed Mpumalanga. HIV education is largely limited to campaigns in schools, clinics and multi-media campaigns which have a significant outreach to rural areas. HIV-related services were significantly strengthened in all the local clinics in advance of the start of the study. Available services include free condom distribution, the practise of syndromic STI management and the provision of voluntary counselling and testing (VCT) for HIV.

4.03 Data Collection

4.03.1 The IMAGE Study

This thesis was researched within the context of a larger ongoing study in South Africa. The IMAGE (Intervention with Microfinance for Aids and Gender Equity) study is a programme of research based in the Sekhukhuneland district of South Africa's Limpopo Province since 2001. The study is a matched-community intervention trial encompassing three years of follow up, centred around eight villages. The study seeks to investigate the potential role of a structural intervention with two components – namely the operation of a poverty-focused microcredit programme targeted at women and a participatory learning and community mobilisation programme for those who join the programme – in alleviating poverty, empowering women, altering attitudes and communication, changing sexual behaviour and preventing HIV infection (203-205).

The study was developed in South Africa, as a collaboration between the University of the Witwatersrand (Johannesburg), the London School of

Hygiene and Tropical Medicine (LSHTM), the South African National Department of Health and Welfare and the Small Enterprise Foundation (SEF), South Africa. The IMAGE Study has been given ethical clearance by ethical committees at both the London School of Hygiene and Tropical Medicine and the University of the Witwatersrand, Johannesburg.

The author of this thesis has been involved with the IMAGE programme since the early design phase in 2000 and since that time has been part of a four person committee responsible for all major decisions regarding the work, supported by experts from the various partner institutions. The author was involved in all of the following processes, leading those marked with (*).

- Study design* (2000 – 2002)
- Establishing a field site * (2000 – 2001)
- Questionnaire design and piloting * (2000 – 2001)
- Fieldworker training * (July 2001 – 2004)
- Management of fieldwork* (2001-2005)
- Database design, management and supervision of data entry and cleaning * (September 2001 – 2005)
- Intervention planning

This PhD thesis was conceptualised by the author, in collaboration with supervisor Linda Morison. All aspects of the design, analysis and write-up of the research described in this thesis have been conducted by the author.

4.03.2 Data sources

Data presented in this thesis come from two main sources. Firstly, participatory wealth ranking was used to generate a sampling frame and to provide data on relative household wealth. Secondly, a household survey was conducted that provided further data on household wealth, as well as individual-level data on socioeconomic roles, HIV awareness, sexual behaviour, HIV infection and potential confounding variables.

4.03.3 Participatory Wealth Ranking

Participatory wealth ranking (PWR) was conducted according to operational guidelines developed by the Small Enterprise Foundation (SEF) (206, 207). The usual operational aim of the PWR process is to identify the poorest households within communities in order to target their inclusion in SEF's microfinance programme. In this study, PWR was also conducted in four villages where no microfinance services were to become available. In both sets of villages, the only incentive to take part in the ranking process was the provision of refreshments (usually soft drinks, bread, peanut butter). All stages of PWR were facilitated by a trained SEF staff member, and data were collated on standardised forms. A period of introduction of the study with village leaders was undergone in every village prior to starting work, and their support for the study was sought and obtained. However, individual informed consent was not obtained for the participatory wealth ranking process as this was not part of the procedures of the microfinance organisation who conducted the rankings.

(a) Mapping and generation of a sample frame

Community members were invited to an open meeting in the village. After introduction of the project, definable "sections" of the village were identified, and the boundaries agreed on. Following this, groups of individuals residing in defined village sections got together to draw a map of their residential area. Typically this area held 50 – 200 households. The participants numbered all households on the map and provided a list of household head names or other dwelling identifiers on numbered cards. This process took one day on average. If a section of the village was identified, but no or very few residents attended the meeting, a repeat visit was made by the facilitator on another day to collect this data.

(b) Wealth ranking

Wealth ranking was generally conducted the day after mapping. Smaller meetings were held with 4-6 residential area members at a time. These were usually predominantly women from poorer households, although any adults from the village section were eligible to participate.

The group was first led in a facilitated discussion on aspects of poverty in the village. Participants were asked by the facilitator to characterise households that were "very poor", those that were "poor, but a bit better off" and those that were "doing OK". These questions were posed in turn to participants and the proceedings of the ensuing discussion were captured by the facilitator in the form of short statements.

Households in a given section were then ranked from the poorest to the most wealthy with reference to the definitions provided. Cards marked with household identifiers were randomly selected and the group asked to compare them with other households in the area. The process started with a comparison of the first two cards selected. As the process proceeded, a number of piles of households deemed by participants to be of similar wealth were generated, with a range from the poorest to the wealthiest. At the end of this process, participants were asked to describe the characteristics of the households in each ranking pile. Each pile was discussed in turn and these discussions were also recorded by the facilitator in the form of short statements. Neither the number of ranking groups nor the number of households that were to be put in each group was determined in advance, although at least four separate ranking piles were required to make the process valid.

The ranking process was then repeated twice more with different groups of 4-6 community members, so that each household was ranked on three separate occasions.

(c) Data Management and Quality Assurance

The validity of maps generated for two of the smaller villages was checked using a rapid village walk procedure over 1-2 days. In one village a small number of houses appeared to have remained unmapped. In the other village the pattern was similar, however, additionally one remote section of the village, comprising approximately 25 households had remained unmapped. This section was later mapped and ranked in a similar way to the other sections. It also became apparent some months after the finalization of the sampling frame that one other village section had remained unmapped in another village.

All data, including household identifier numbers, rankings, and the short statements made during the process were entered into purpose-built Microsoft Access databases. Data were cleaned through running logical queries in the database to identify outlier values, and referring to source data to correct any inconsistencies.

As described in the previous section, data in the form of text statements were collected at two stages of each ranking process. The first of these was before wealth ranking was performed, in response to three general questions about characteristics of households in different wealth bands ("very poor", "poor, but a bit better off" and "doing OK"). These descriptions will be referred to as *general statements*. The second collection of text data occurred after the wealth ranking process was completed, when respondents were asked to describe the characteristics of households in each ranking pile. These statements will be referred to as *pile statements*. The same statements were often made during both of these processes. *General statements* and *pile statements* were coded using the same technique. The coding scheme was devised as coding progressed. Codes were grouped into themes and specific statement codes. Statement codes were kept as specific as possible. An example is the following; "Have food", "Able to buy food", and "Able only to buy food" were each coded under the *Theme* "Food" but were each given separate *statement codes* since each has a slightly different meaning.

Conversely, the statements "no food", "they have no food" and "no food available" were all given the same *statement code* since there is no discernible difference between their meaning. Statements of both types could be assigned up to three separate codes.

4.03.4 Household survey of individuals aged 14-25 years

(a) Sampling Strategy

The eight villages mapped and included in the study were purposively selected to be included in the study on the basis of their size and accessibility. Using the sampling frame generated during the PWR process, 200 dwellings were randomly selected from each of the villages using a procedure in *Stata* (Statacorp, Texas). Selected dwellings were visited at least three times to generate a full household roster. For the purposes of the study a household was defined as a group of people who were residents at the same dwelling, who may have been either staying at home or away from the home at the time of the survey, but who ate from the same pot of food when staying there (28). This definition is in line with other population research programmes in South Africa and was intended to account for the high levels of temporary migrancy in the country and the fact that temporary migrants often maintain significant links with the rural home (12). In most cases, the physical dwellings identified during the PWR mapping process coincided with the residence of a single household. However, occasionally this was not the case, and the fieldworker would then identify all of the households at the dwelling and these were all eligible for inclusion in the survey. Interviewers received extensive training in identifying households, and were given specific examples that attempted to recognise both local norms and the research requirements. For example, the daughter of the head of a household who had married and gone to live with her husband permanently, but who occasionally visited and provided some resources to the household was not considered a household member. In contrast, the daughter of a household head who was staying on a local farm to take up work, and who came home only at the ends of each month, but who did not have any other place that she called home was considered a household member. Inevitably, there were a small number of cases that did not easily fit into such classifications and for which interviewers were required to make judgements based on the core principles.

For each household the household head was also identified. The household head was that person identified by other adults in the household as the household head. As reported for other studies, there were rarely reports of confusion with this definition (28). In general, in the patriarchal society where this study took place, the head of the household will likely be the eldest male in the home, though there are many situations in which this is not the case due to influences such as the separation of married couples, widowing, households taking in elder relatives and others.

All 14-35 year olds recorded on the household roster were eligible for inclusion in the main study, irrespective of whether they were currently sleeping at the dwelling or not. At least three repeat visits were also made to interview these individuals, and an additional period of fieldwork was conducted during the Christmas and New Year period 2001/2, after the end of the main fieldwork period, in order to try and maximize the number of interviews conducted with migrants. The analysis in this study concentrates only on those individuals aged 14-25 years.

(b) Data collection tools

Two questionnaires were used in the study and a biological (oral fluid) sample was collected for analysis in a laboratory. Copies of the data collection tools are held in Appendix 1 and 2. Questionnaires were extensively piloted in a village not included in the final study sample during late 2000-mid 2001. Questionnaires were translated from English to Sepedi by a local speaker involved in the data management process. Questionnaires were then back-translated and differences in interpretation resolved.

The *household questionnaire* was usually administered to a single member of the household, but occasionally to two or more individuals together. It was desirable that either the household head or his / her partner be involved in the interview, although there were many occasions on which this was not possible. The household questionnaire comprised two sections. Firstly, a household roster was compiled and sociodemographic information was collected on household members' residence status, year of birth, marital status, relationship to the household head, education, employment and receipt of state benefits or other non-work related incomes. Secondly, information on the household unit as a whole was collected, including data on the construction and services available to the dwelling, ownership of a short list of assets, credit and savings, food security and experience of economic shocks during the past year. Informed consent for the household interview was obtained from the person interviewed, who was also informed about the intention to interview young people.

The second data collection tool used was known as the *young person's questionnaire*. This was administered to individuals aged 14-35 from the household roster. Verbal informed consent was obtained for the interview which was witnessed by the interviewer who signed that this had been given. This questionnaire was always conducted in a quiet place chosen by the respondent. Respondents were assured of the confidentiality of their responses. If consent for the interview process was given, the interviewer signed the interview sheet to confirm that this had been received. Fieldworkers received extensive training in interview techniques designed to maximise honest reporting. The questionnaire collected further sociodemographic information on the individual (including date of birth and orphan status), and data on communication about sexuality issues, contraception, circumcision (males only), lifetime sexual behaviour, details of spousal and non-spousal partners in the last year, beliefs about gender norms, knowledge about HIV and HIV-testing services and attitudes to individuals with HIV infection.

Finally, participants in the young person's questionnaire were also asked to provide an oral fluid specimen for HIV testing. A separate witnessed verbal consent procedure was used in this regard. Some individuals completed the questionnaire but did not provide a biological specimen. Oral fluid samples for the assessment of HIV status were collected using the OraSure collection device according to manufacturer's guidelines. Briefly, a pad was held in the mouth, between gum and cheek, for at least two minutes before being sealed in a vial. Specimens were stored at room temperature in the field and then transported to the National Health Laboratory Service in Johannesburg.

Qualitative determination of antibodies to Human Immunodeficiency Virus (HIV) type 1 and/ or 2 (anti-HIV-1, anti-HIV-2 and anti-HIV-1 group O) from collected samples was performed with the Vironostika HIV Uni-Form II assay. The OraSure specimen vial was centrifuged and specimen eluted from the collection pad. The eluate was diluted and added to the microelisa wells. Each microelisa well contained a horseradish peroxidase-labelled conjugate sphere of the same HIV-antigen mixture. Test samples were then incubated in the microelisa wells. With the presence of antibodies to HIV-1, HIV-2 or HIV-1-group 0 a solid phase antigen/anti-HIV/enzyme labelled antigen complex formed. Following a wash procedure and incubation with substrate, colour develops which turned yellow when the reaction was stopped with sulfuric acid. If anti-HIV-1, anti-HIV-2 and/or anti-HIV-1 group 0 was present in the sample, an intense colour developed. However, when the sample was free of anti-HIV-1, no or low colour formed with the addition of substrate.

The result of analysis of the oral fluid sample was not offered to participants. However, all participants were referred to local health services where HIV testing alongside full pre- and post-test counselling were available.

(c) Data Management and Quality Assurance

Data collection was conducted by female fieldworkers recruited from the study villages. Fieldworkers did not conduct interviews in their own village of residence since it was felt this may prove problematic given the sensitive nature of some of the data to be collected. A team of fifteen fieldworkers was recruited. The team was overseen by a fieldwork supervisor, and support was provided to fieldwork in the form of overall direction, technical support and quality assurance from the author of this thesis, with assistance from a volunteer. Fieldworkers received an intensive four week training period in general interviewing skills and the specific details of the data collection tools, and were required to pass an assessment before commencing work. Fieldwork was guided by a manual that documented all aspects of questionnaire conduct, recording of work completed and field procedures.

Data on sexual behaviour were collected during the middle of the interview, after rapport had been established between interviewer and respondent. Before asking questions on sexual behaviour, the interviewer briefly stopped the interview and re-iterated to the respondent the sensitive nature of the issues about to be discussed. The right of respondents to refuse to answer any question at any time was re-asserted, alongside a reminder of the importance of answering questions honestly if an answer was given.

A high level of field supervision was maintained during the fieldwork period. Fieldworkers were required to have a personal meeting with the supervisor every week. Occasional supervised interviews were conducted with all interviewers during the fieldwork period and feedback given. Questionnaires were checked at least three times before data entry, with illogical data being queried and returned to the fieldworker responsible. Occasionally this necessitated revisiting the respondent, but more often required clarification on the questionnaire without revisit. This process served both to identify problems after they had occurred and to ensure a high quality of fieldwork.

Data from the questionnaires were then entered directly into purpose designed Microsoft Access databases. An intensive data cleaning process was implemented through the application of logical queries to check for outliers, illogical variable arrays and missing data. Inconsistencies were checked against source data and corrected where possible. Following this, a subset of questionnaires were re-entered to the database and over 99.5% accuracy in data entry was noted.

Laboratory data were entered at the laboratory into Microsoft Excel spreadsheets and then transported to the central field office for data cleaning. Data were imported to Microsoft Access and logical queries to identify missing and outlier data run, as well as cross checking with questionnaire data. Problems within the data were solved by checking the source data. Source data that were illogical or missing were then coded as such in the database.

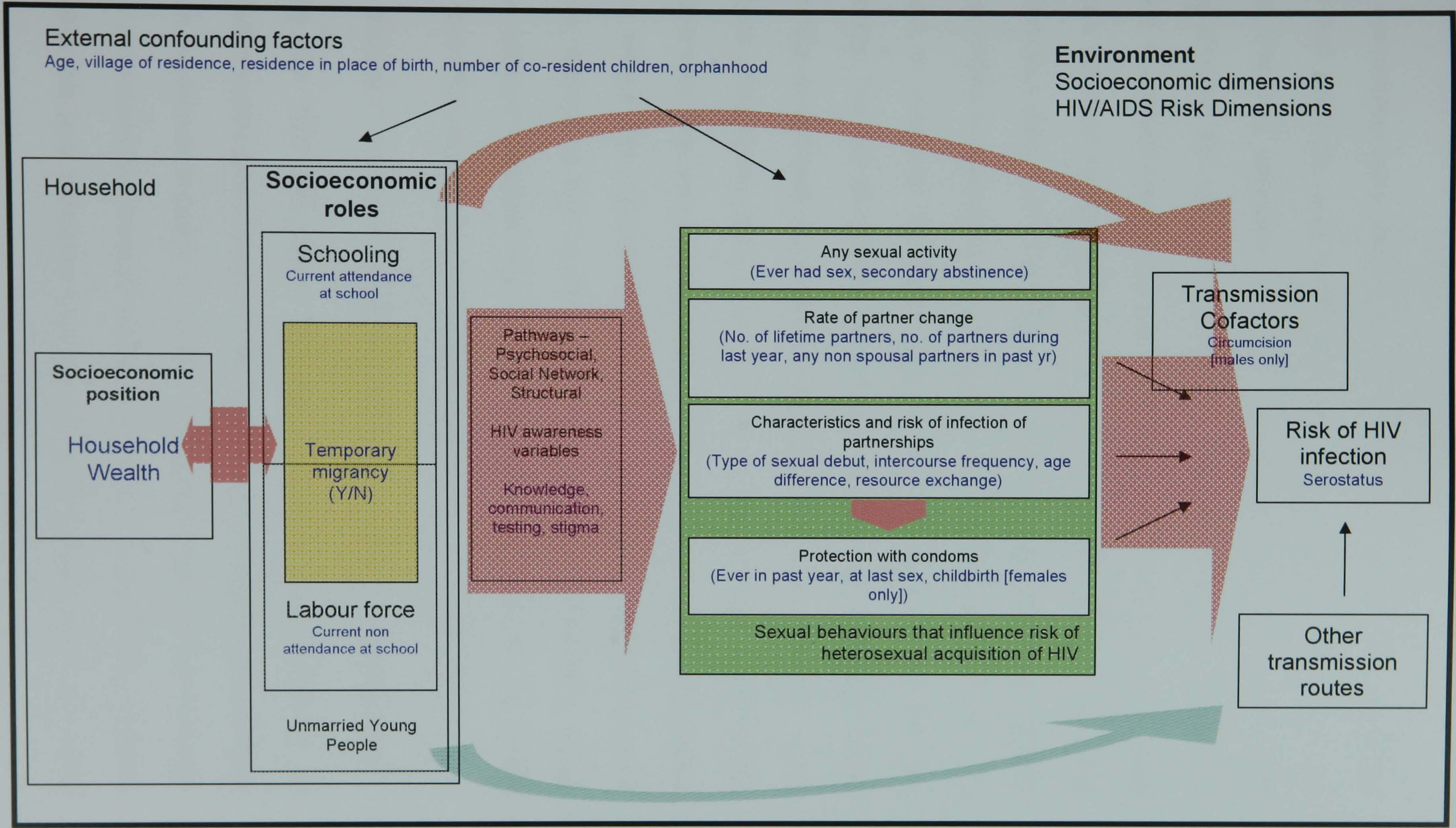
Prior to data analysis all data were "de-linked" from the hard copy data through the removal of any text or unique identifier variables. A first set of descriptive analyses were performed with field staff involved in the collection of data. Distributions of key variables were tabulated and field staff were asked to reflect on issues in asking these questions in the field. These discussions were documented and reviewed. This led to the identification of some variables that had been differentially or incorrectly asked by some fieldworkers.

4.04 Operationalising the study framework

Figure 4.2 shows the conceptual framework developed in Chapter 2 with details of the variables operationalised in the current study included. Red arrows indicate relationships that are directly under study within this thesis, while the green arrow indicates a pathway that is not directly under study in this thesis due to data limitations. The boxes presented in the conceptual framework diagram depict important underlying constructs and concepts. Blue text depicts actual variables operationalised as proxies to mark for the underlying concepts. Important factors on which no data were available, such as prevalent STIs, are not presented in the framework diagram but are discussed later in the thesis.

The remainder of the chapter provides details of the way in which variables have been operationalised in the current study.

Figure 4.2: Schematic diagram depicting key components of the conceptual framework with variables operationalised



4.04.1 Household wealth

Three approaches were used to generate a measure of relative household wealth that could be used to rank households from the wealthiest to the poorest. The first approach used data only from the participatory wealth ranking process; the second used data on indicators of household welfare from the household survey, but with the selection and weighting of these indicators informed by the wealth ranking process; the third approach used only indicator data from the survey, employing a principal components analysis (PCA) to determine the weights. While the approaches were each designed to measure household wealth, they differed in at least two dimensions. The first dimension was whether information on aspects of the household's wealth were provided by household members (as for both of the techniques using survey data), or by other community members (for the PWR only approach). The second dimension was whether community views were used to weight the importance of different aspects of wealth (an *emic* approach, intrinsic in both of the approaches that used the PWR data), or whether external statistical rules were used (an *etic* approach, as in the survey data PCA method used here). Each approach is described below.

(a) Method 1 : An index of household wealth from participatory wealth ranking

Following the field procedures described above, it was necessary to manipulate the data obtained from PWR to generate an indicator of household wealth. A traditional limitation of wealth ranking data are that the rankings provided are intrinsically linked to the ranking session in which they occurred. Households are only compared directly with other households in the same village section. Additionally, although each household is ranked against its neighbours three times, a household ranked in the "wealthiest" of four groups in one ranking session might be ranked in the "second wealthiest" of six groups in another session, even though the participants agree on the wealth of the household. To overcome this limitation a novel approach to the use of the wealth ranking data was developed (208). This method sought to link the index developed from the ranking data not only to the relative rank a household was given compared to its neighbours, but also to qualitative statements made about the absolute and relative welfare level of households grouped together. The following steps were taken.

1) Assigning a score to the ranking piles

Households within each village section were ranked by three independent groups of PWR participants. Within each of these three ranking processes a number of "ranking piles" of households were generated. Each pile was assigned a score such that the wealthiest pile (pile 1) received a score of 100 and the poorest pile (pile N) received a score of 0. The scores for the remaining piles were calculated as *Score for pile n = 100*((N-n)/(N-1))*, where n was the pile number and N was the total number of ranking piles. For example, in a situation where five ranking piles were formed (N=5), pile 1 (wealthiest) was given a score of 100, the next pile a score of 75, then 50, 25 and finally 0 for pile 5 (the poorest pile). It should be noted that the scoring system used in this thesis is a slight variation on the pile scoring system used in operational work by the Small Enterprise Foundation.

2) Pile statement scoring

Coded *pile statements* were made in relation to each of the ranking piles generated during the three ranking processes conducted in each village section. These were assigned the numeric score allocated to the pile (as described above). An average score was then calculated for each *pile statement*. The average *pile statement* score was calculated as the mean of the pile scores to which that statement was associated, covering the full PWR process in all eight villages. For example, suppose the statement "able only to buy food" was made 10 times during all the ranking sessions that occurred. Suppose further that the statement was made eight times in relation to the poorest pile in given rankings (scoring 0), and twice in relation to piles that were ranked second poorest in ranking sessions that generated five wealth ranks (scoring 25 each time). In this case, the average pile statement score for this statement would be ((0*8)+(25*2))/(10) = 5. A *pile statement* score was generated for all statements made more than three times during the entire PWR process covering all villages.

3) Generating a household wealth index

As a result of the three ranking processes for each village section, each household was included in three piles. Each of these three piles had *pile statements* associated with it, and each pile statement had been allocated an average *pile statement* score (as described above). A household wealth index for each household was then calculated as the mean of the *pile statement* scores of all the pile statements made in relation to the three piles into which that household was ranked. This technique ensured that there was a direct link between the household wealth index and the score each statement received. Households, like statements, could receive a theoretical maximum score of 100 and a minimum score of 0.

4) Generating poverty lines

Uniquely of the three wealth indices, this methodology allowed households to be separated into groups in a manner directly informed by the data. In order to generate these "poverty lines", a list of all the pile statements was made in ascending order of their average *pile statement* scores (see Appendix 3, or table 6.1 for an abridged version). Alongside each *pile statement* was included a count of the number of occasions each statement was mentioned in the three *general statement* categories "very poor", "poor but a bit better off" and "doing OK". A visual assessment was then used to see if obvious cut-off scores could be applied to group the *pile statement* scores into categories corresponding to "very poor", "poor, but a bit better off" and "doing OK".

As the household wealth index was composed of an average of *pile statement* scores it was then judged appropriate to apply the same numeric cut-off scores identified in the visual assessment to the household wealth index. The final stage of the analysis was to apply these cut-off scores to the household wealth index to group households into wealth bands on the basis of these local perceptions of poverty.

(b) Method 2 : An index of household wealth developed from household survey data but informed by participatory wealth ranking

The second measure of household wealth combined information from the survey and the PWR process. The data used to generate an indicator of household wealth came from the survey. However, the decision of which factors to use in developing this composite indicator of household wealth, and how to weight these data was informed by the results of the PWR process. Table 4.3 lists the range of "themes" held in statements made by PWR participants when asked to describe aspects of wealth in their village. It was decided that where data was available on aspects of household wealth relating to each of the 10 commonest themes, this would be used to inform the calculation of the index of household wealth.

Having decided on these themes, it was then necessary to decide which data from the survey to use and how to weight the data that was provided on these themes. The data available and the way in which it was pre-coded was limited by the fact that the data to be collected in the survey was decided on before the PWR process was complete. The scoring applied to the survey data to create the index was informed by the scoring for specific statements made in relation to each theme from the PWR process. For each theme a maximum score of 2 (relating to a higher level of wealth) was theoretically available, and a minimum score of -2, although for some themes the maximum and minimum scores were not possible as data from the survey did not link directly to the statements in a way that allowed this. Table 6.3 holds final details of the scoring system applied. On the basis of this scoring system, each household could receive a maximum score of 9 (wealthiest) and a minimum score of –10 (poorest).

Table 4.3: Themes raised in discussions of poverty during participatory wealth ranking conducted in rural South Africa

Theme	*General Statements*		*Pile Statements*		*Total*	
	N	%*	N	%*	N	%*
Employment	577	14.5%	2236	21.3%	2813	19.4%
Schooling	557	14.0%	1351	12.9%	1908	13.2%
Housing	559	14.0%	1229	11.7%	1788	12.3%
Food	506	12.7%	921	8.8%	1427	9.9%
Self employment	380	9.5%	922	8.8%	1302	9.0%
Clothing	399	10.0%	755	7.2%	1154	8.0%
Family and household	146	3.7%	781	7.4%	927	6.4%
Money	162	4.1%	535	5.1%	697	4.8%
Pensions	100	2.5%	499	4.8%	599	4.1%
Cars	177	4.4%	297	2.8%	474	3.3%
Begging	91	2.3%	134	1.3%	225	1.6%
Towns	0	0.0%	216	2.1%	216	1.5%
Need	10	0.3%	96	0.9%	106	0.7%
Dirtiness	42	1.1%	60	0.6%	102	0.7%
Water	28	0.7%	49	0.5%	77	0.5%
Survival skills	11	0.3%	65	0.6%	76	0.5%
Societies and stokvels	34	0.9%	37	0.4%	71	0.5%
Other assets	34	0.9%	35	0.3%	69	0.5%
Health	31	0.8%	31	0.3%	62	0.4%
Furniture	24	0.6%	24	0.2%	48	0.3%
Livestock	19	0.5%	26	0.2%	45	0.3%
Planning	16	0.4%	25	0.2%	41	0.3%
Fighting	6	0.2%	29	0.3%	35	0.2%
Happiness	16	0.4%	19	0.2%	35	0.2%
Depended upon / employ ...	20	0.5%	14	0.1%	34	0.2%
Telephones	8	0.2%	25	0.2%	33	0.2%
Employ home help	3	0.1%	23	0.2%	26	0.2%
Position within society	3	0.1%	20	0.2%	23	0.2%
Electricity	11	0.3%	11	0.1%	22	0.2%
Crime	10	0.3%	5	0.1%	15	0.1%
Whites	3	0.1%	11	0.1%	14	0.1%
Land / Agriculture	9	0.2%	3	0.1%	12	0.1%
Lifestyle	0	0.0%	9	0.1%	9	0.1%

* Reported percentages are as a percentage of all coded statements

(c) Method 3 : An index of household wealth developed from household survey data, with weightings informed by principal components analysis

To generate the third measure of household wealth, variables to be included in the questionnaire that captured salient aspects of household wealth were decided upon by the researcher following literature review and extensive piloting in the local area. It was decided in advance to make all of these variables available for entry into the principal components analysis. Fourteen variables were created from the survey data under four subheadings; value of selected household assets per person, quality of housing conditions, income/human capital and food security. These variables are shown in more detail in Table 4.4. The value of assets was derived from the survey data by multiplying the number of owned assets of each type that was new (less than 2 years), relatively new (2-6 years), or old (>6 years) by estimations of the value of those assets which came from a small sub-study in which participants were asked to estimate the average re-sale and purchase values for the items listed. The other variables were drawn directly from the questionnaire. By including variables on such a wide conceptualisation of welfare, the technique used was similar to that developed by the Consultative Group to Assist the Poorest (CGAP) (93). Non-continuous variables were coded even-spaced ordinally. The association between categorised versions of each of the 14 variables against each of the others was examined. The value of selected non-livestock assets per person was selected as a key marker of household wealth. Factors not significantly associated in the expected direction with this in a Chi-square test ($p<0.05$) were not included in the later analysis.

Table 4.4: Survey variables entered to principal components analysis to develop an indicator of household wealth

Variable	Coding	Mean (SD), Range of variable entered to PCA
Assets		
Estimated value of selected non-livestock assets per person	-	1550.9 (3225), 0-76664
Estimated value of selected livestock assets per person	-	874.5 (1813), 0-21860
Land tenure	0 = No, 1= Yes	0.3 (0.4), 0-1
Housing		
Quality of house wall material	0=Tin, 1=Mud and sticks, 2= Mud bricks without cement, 3=Mud bricks cement covered, 4=Block bricks without cement, 5=Block bricks cement covered, 6=Face bricks	3.9 (1.5), 0-6
Quality of toilet facility	1=No facility, 2= Basic, 3=Modern	1.8 (0.4), 1-3
Household Electricity	0=No, 1=Yes	0.7 (0.5) 0-1
Accessibility of water supply	1=Low, 2= Medium, 3=Good	1.7 (0.5), 1-3
Density of household living conditions (rooms per person)	-	0.9 (0.8), 0.1-8
Income / human capital		
Proportion of household members receiving a regular income	-	0.2 (0.2), 0-1
Educational level of household head	1=No formal schooling, illiterate, 2=No formal schooling, literate, 3=Some primary, 4=Completed primary, 5=Some secondary, 6=Completed secondary, 7=Attended technical college / vocational college, 8=Attended University	3.0 (1.7), 1-8
Percentage of household members working age adults	-	0.6 (0.2), 0-1
Gender of household head	0=Female, 1=Male	0.6 (0.5), 0-1
Second most important household income	0=Non-Financial, 1=Financial	0.6 (0.5), 0-1
Food security		
Regularity of household having a meal consisting of mielie meal alone, bread alone or worse	1=Often, 2=A few times, 3=Once only, 4=Never	2.3 (1.2), 0-4

Non-livestock assets comprised cars, televisions, hi-fis, fridges, bicycles, cellphones. Livestock assets were cows, goats, chickens. Low accessibility of water supply was defined as those collecting rain or stream water, medium level access was through a borehole or tap in the village, while those with high quality access were those with a tap in the plot of the dwelling.

The remaining factors were included in a principal components analysis (PCA). PCA is a statistical methodology that transforms a set of correlated variables into a set of uncorrelated 'components'. When a set of variables that each hold information about some underlying concept are all included in the analysis, PCA can be used to produce the best single composite variable among all possible linear functions of the original variables. This composite variable, which is the component explaining the greatest proportion of the total variance, is called the first principal component and implicitly weights the data used to generate it in proportion to how well each variable is correlated with the others. Since all factors entered into the model were selected to indicate some aspect of household welfare, the first principal component can be used as a single indicator of household wealth.

PCA was run using Stata v.8. The correlation matrix, as opposed to the co-variance matrix, was used in the analysis since each of the different variables had different variances and it was not required that the variance of each variable should affect its weighting in the first principal component. An Eigen value was calculated for each component extracted from the data, and the component data added to the model are each assigned a scoring coefficient denoting the relative weight applied to each factor in the final analysis. Each PCA model was run using only the available data, so first principal component score could only be calculated for households where data was available on all the variables.

A number of models were run, with factors being excluded according to the following rules; factors with the lowest component loadings on the first principal component were excluded since these were not well correlated with the other variables. A component loading factor of less than 0.2 was used as a guideline for these exclusions. The remaining nine factors were included in the final model. The first principal component was calculated for each household, and across the full sample this component was approximately normally distributed and had a mean of approximately 0 and a standard deviation of 1.

4.04.2 Socioeconomic roles

Two variables relating to socioeconomic roles played by young people were operationalised; firstly, attendance at school, and secondly, temporary migrancy.

Data on the main socioeconomic activity undertaken during the previous year was drawn from two questions in the household questionnaire and used to construct a variable on school attendance. One question asked, "Are you currently in school?", the second asked about the primary function of the individual over the course of the past year. Individuals in school were marked as "Student" for their primary function during the past year. In the very small number of cases (<5) where the two variables suggested different things, the data directly enquiring about current school attendance were used. Those marked as "in school" / "student" included those in primary, secondary, or full-time further education. This did not include individuals attending adult education training or part time or correspondence higher education courses which are quite common in South Africa. Among those who were not students, a wide variety of codes were available to describe their main activity during the past year. These codes were "Self employed in agriculture", "Self employed in non-farm enterprise – registered business", "Self employed in non-farm enterprise – unregistered business", "Salaried worker", "Domestic worker", "Unemployed, looking for a job, often does casual, seasonal or contract work", "Unemployed, looking for a job, occasionally gets any casual, seasonal or contract work", "Unemployed, looking for a job, rarely or never had any work during the last year", "Unwilling to work, retired or too young to be working", "Unable to work (Handicap)". As will be seen in Chapter 6, the number of individuals in registered businesses and salaried employment was low (<3% of the whole sample). For the purposes of examining the impact of attendance at an educational institution on HIV risk outcomes, individuals in all of these categories were grouped together since it was the potential impact of attendance at school on access to HIV prevention campaigns or sexual network structure that was the primary reason for interest in this variable.

Data on migration came from a combination of two questions, "Is X currently sleeping at home in the last month?", and "For how many months of the last year were you staying here?". The first question enquired of the current status of each household member as to whether they were currently (in the month prior to the interview) mainly sleeping at the household dwelling or away from the household dwelling. The second question was asked of young people only and enquired about the whole of the past year. Individuals may have been sleeping at the family home during all of the past year (12 months), may have been living away from home for most of the year and visiting home most weekends (approximately 3 months), or month ends only (1 month), or other patterns. For the purposes of statistical analyses a binary variable was created, "temporary migrant", which was coded 'Yes' for individuals who were either not currently sleeping in the home OR reported that they had not been sleeping in the home for more than six months of the previous year.

4.04.3 HIV risk outcome variables : HIV awareness, sexual behaviour and HIV serostatus

Table 5.4 gives details of the questionnaire items used to develop outcome variables and, where relevant, the way in which data from more than one question were combined to create variables relating to these issues. Each outcome variable was coded into binary form. In many cases this was the natural form of the variable since questions required a Yes / No answer. However, in some cases continuous or ordered categorical variables were reduced to binary variables using cut-offs.

In each case, the binary outcome variable was coded '0' for the state associated in previous literature with low potential risk of HIV infection, while the state associated in literature with high potential risk of HIV infection was coded '1'. "High risk" outcomes included low knowledge of HIV, lack of communication with parents, lack of HIV testing, holding stigmatising attitudes, being sexually active, not wanting to have sex at sexual debut, having multiple sexual partners, having had a child, lack of circumcision, more frequent sexual intercourse, male seniority in age (for females) or lack of male seniority (for males), economic asymmetry (male providing resources, female not), lack of condom use and being HIV positive. For three of these variables some additional explanation of the coding system is necessary. One variable asked about experience at sexual debut. While previous literature particularly suggests that violent sex at debut may be associated with greater HIV risk, this response was grouped together with those who said their sexual debut "just happened". The reasons for this was that there was different interpretation between fieldworkers as to what constituted being "forced" to have sex at debut, while there was greater agreement on whether individuals "wanted" to have sex or not. Regarding age difference within partnerships, literature suggests that a greater male-female seniority is a risk factor for infection of the female partner (50, 209). For males, most partners are younger and so it is partners of a similar age who have had a longer time to become infected and are thus higher risk. Finally, resource exchange was commonly reported in relationships. Following the literature, partnerships where resource exchange patterns suggested that the male partner may be in a stronger bargaining position with regard the timing and characteristics of sexual intercourse (defined as those in which the male partner was providing resources but the female partner was not) were coded as those of higher risk.

Table 4.5: Coding of binary HIV risk outcome variables including HIV awareness, sexual behaviour and HIV serostatus

Variable : HIV Risk Characteristic	Original questions (see also Appendices)	Coding for outcome analysis	
		0 "Low risk"	1 "High risk"
HIV awareness		If ...	If ...
Low HIV Knowledge	(i) Do you think that a healthy-looking person can be infected with HIV, the virus that causes AIDS? (ii) Can a pregnant woman infected with HIV transmit the virus to her unborn child?	All questions answered "Yes"	Any question answered "No" or "Don't know"
Lack of household communication	(i) Have you spoken about sex, and sexuality in general to your parents / guardians during the past 12 months? (ii) Have you spoken about sex, and sexuality in general to other household members during the past 12 months?	Any question answered "Yes"	All questions answered "No"
Lack of HIV testing	I don't want to know the result, but have *you* ever had an HIV test?	"Yes"	"No"
Holds stigmatising attitudes	(i) Would you be willing to share a meal with a person you knew had HIV or AIDS? (ii) If a relative of yours became ill with HIV, the virus that causes AIDS, would you be willing to care for him in your household? (iii) If you knew a shopkeeper or food seller had the HIV virus, would you buy food from them?	All questions answered "Yes"	Any question answered "No" or "Don't know"
Individual sexual behaviour characteristics / co-factors			
Sexual debut	(i) Have you *ever* had sexual intercourse? (ii) At what age did you first have sexual intercourse? *	"No"	"Yes"
Type of sexual debut#	How would you describe the first time that you had sex? Would you say that you wanted to have sex, you did not want to have sex but it happened anyway, or were you forced to have sex?	"Wanted to have sex"	"Did not want but just happened" or "Forced to have sex"
Multiple partners (lifetime)	How many people would you say you have had sexual intercourse with in total up to now in your life?	Males: <=3 Females: <=2	Males:>3 Females:>2
More than 1 partner (past year)	How many of your partners in the last 12 months were sexual partners that you are not married to and have never lived with?	0 or 1	> 1
Lack of circumcision [males only]	Have you been circumcised?	"Yes"	"No"
Childbirth [females only]	How many children have you had up to now in your life?	0	>=1
Continuing sexual activity (opposite "secondary abstinence") #	(i) Have you *ever* had sexual intercourse? (ii) Have you had sexual intercourse in the last 12 months?	(i) "Yes" (ii) "No"	(i) "Yes" (ii) "Yes"
Partnership Characteristics			
Frequent sexual intercourse ^	During the last 12 months how often would you say you have had sexual intercourse with this person?	"Once only" or "2-5 times"	"6-20 times" or ">20 times"
High partner's age^	[Respondent's age] calculated as Date of interview – Date of birth, rounded down to nearest year. [Age of partner] How old is that person (the non-spousal partner)?	Males: *(Respondents Age) – (Age of partner)* >3 Females: *(Respondents Age) – (Age of partner)* <=3	Males: *(Respondents Age) – (Age of partner)* <=3 Females: *(Respondents Age) – (Age of partner)* >3
Male partner only provides resources^	(i) Do you regularly provide financial support to this person ? (ii) Do you regularly receive financial support from this person ?	Males: (i) "Yes" and (ii) "Yes", or, (ii) "No" Females: (i) "Yes", or, (i) "No" and (ii) "No"	Males: (i) "Yes" and (ii) "No" Females: (i) No and (ii) "Yes
Never used condom in past year^	How often would you say you have used a condom when having sex with this person in the last 12 months?	"Less than half the times" or "Half or >half the times" or "Always or nearly always"	"Never"
No condom use at last sex^	Did you use a condom the last time you had sex with this person?	"Yes"	No
HIV positive	-	HIV negative	HIV positive

* For analysis of this variable, survival techniques were used with information censored at current age or age of sexual debut
# N/A if never sexually active
^ N/A if no sexual partners during the last year

Finally, data was available for male participants on male circumcision and, where data was available, HIV serostatus was operationalised as positive or negative on the basis of lab procedures described above.

4.04.4 Confounding variables

Five potential confounding variables were considered throughout the analysis – age, village of residence, still residing in place of birth, number of other young people aged 14-25 years in the household and orphanhood status. For females, ever having had a child was additionally considered as a potential confounder. In each case these were factors that were theorised potentially to be associated with at least one or both of the exposure and outcome variables, but not to lie on the causal pathway between them. In some cases this could be disputed, and this is an issue discussed in greater detail in Chapter 10.

Age was calculated as the age in years from date of birth to date of interview and was categorised into either two or four categories dependent on the analysis. Village of current residence was coded from 1 to 8, and was entered as a categorical variable into regression models. In a small number of analyses it was necessary to group villages into the four matched pairs employed for the main study, if no outcome events occurred in a single village. Place of birth was categorised for use in analyses as those who were still currently living in the place where they were born compared to those who were not. The number of other 14-25 year olds of either sex within the household was used as a proxy for competition for resources, and was coded '0' for those who were either the only young person or there was only one other, compared to those who had more than one other co-resident young person (coded '1'). This might confound the association between socioeconomic factors and HIV risk outcomes if co-resident young people are in competition for household resources to attend school or migrate or are from larger households that are more wealthy, and if co-residence of other young people affects HIV awareness (for example, through increased intra-household communication) and consequently sexual behaviours. Orphanhood was coded to a single binary variable on the basis of two questions relating to the mother and father of the young person. Individuals whose mother or father was either known to be dead or whose status was unknown were coded as orphans. Orphans may come from poorer households, may be less likely to attend school or migrate and may have more risky sexual behaviour profiles. Consequently orphanhood may confound the association under study. This simple identification of orphans without regard for which or how many parents are alive may be a simplification since this may affect the impact of orphanhood, but was felt suitable for the purposes of adjustment..

Finally, ever having had a child was coded as described in the previous section.

4.05 Statistical analysis

4.05.1 Description of the population

Chapter 5 describes the data obtained in the study. The chapter has three sections. Firstly, a general description of the volume of data collected by the study procedures is presented in the forms of numbers of households and individuals mapped, enumerated and interviewed. The second section gives reasons why some individuals eligible for inclusion in the study were not included in the final study sample, and goes on to compare available sociodemographic characteristics between those who were and were not included. Numbers and percentages are presented in this regard, with Chi-square tests used to generate statistical tests of the significance of any differences noted. Finally, the third section of the chapter describes in more detail the sociodemographic profile of the households and individuals that made up the study sample.

Individuals were eligible for inclusion in the analyses presented in the remaining chapters if they were aged 14-25 years on the date of interview. Individuals who had ever been married were excluded from the analysis. The only other reason individuals were not included in analyses was if data was not available. For individuals to be included they needed to have data available on all the major variables. This included data on household wealth, both measures of individual socioeconomic role, four HIV awareness variables, ten sexual behaviour variables for men and additionally childbearing for women, circumcision among men, and five potential confounders. The only exceptions to this rule were in the variables denoting age at first sex, age difference between sexual partners and HIV serostatus. For a subset of individuals who reported being sexually active, age at first sex was unknown. Thus, descriptive data on proportions of individuals sexually active report data on the full study sample, while survival analyses relating to the age at sexual debut have some missing records detailed in the text. Since there was a large amount of missing data ("unknown") was present relating to the age of sexual partners, individuals were not excluded from the analysis if this information was not available. Levels of missing data are described in the relevant places throughout the text. Additionally, some individuals provided data on all other major factors, but did not provide a sample for HIV testing, and these individuals were also considered in the thesis. Where HIV was the outcome, analyses relate to a slightly different number of individuals, and this is highlighted throughout in the text. The small amount of missing data in other less important fields is presented descriptively and noted in the text.

All analyses presented in chapters 6-8 were presented stratified by sex. Descriptive analyses held in chapters 6 and 7 were also stratified by age to provide more detail on the age-profile of important socioeconomic or HIV risk characteristics.

4.05.2 Socioeconomic factors

Chapter 6 provides detailed information on the distribution of key socioeconomic variables developed for the study by age and sex, and examines their associations. The first section concentrates on measures of household wealth. The distribution of key indicators of household welfare from both the PWR process and the survey are presented. Following this, the frequency distribution of each of the three indices of household wealth across households is also shown. For the PWR method (Method 1) it was also possible to provide two additional pieces of information. Firstly, as described above, it was possible to group households into categories of relative wealth directly on the basis of the data provided. Secondly, it was possible to generate a measure of the internal consistency of the PWR process since each household was ranked by three separate groups of PWR participants. For this analysis each household was assigned the pile score for each of the three rankings. An estimation of the intraclass correlation coefficient for these three scores gives an estimate of the consistency with which the three rankings agree with each other on average. A random-effects two way ANOVA model was used to generate this ICC and its 95% confidence intervals using SPSS version 11 (210). Using this approach, the more similar the observations, the higher is the value of the coefficient. A value of 1 for an ICC denotes perfect agreement, and it is probably safe to suggest that values above 0.75 can be interpreted as indicating good consistency, although such a cut off is arbitrary (211).

Following description of the household wealth indices generated using the three different approaches, it was necessary to assess the association between the techniques. Two statistical methods were used for this purpose. Firstly, a kappa value was calculated to ascertain the degree of *agreement* between the methods in assigning households to wealth-rank terciles. For this measure, a value of 1 denotes complete agreement, while 0 represents no agreement. As for the ICC measure discussed earlier, it is probably safe to assume that values above 0.75 can be interpreted as indicating good agreement. Secondly, the Spearman rank correlation coefficient was used to assess the strength of the *correlation* between the indices in their continuous data form.

The remainder of Chapter 6 presents data on the socioeconomic roles of young people in the study sample. The distribution of these variables is presented grouped by age, separately for each sex. The association between attendance at school and temporary migrancy was assessed using a chi-square test with Mantel Haenszel methods used to adjust for age. Finally, the association between each socioeconomic role and household wealth is presented in the form of pie charts, and the strength of the association was assessed using an age adjusted Mantel-Haenszel Chi-square test.

4.05.3 HIV awareness, sexual behaviour and HIV infection

Chapter 7 provides detailed information on the distribution of outcome variables, including HIV awareness variables, sexual behaviour and prevalent HIV infection. Data are presented in four sections. Firstly, descriptive data are presented on HIV awareness factors, including knowledge, communication with parents, HIV testing and stigmatising attitudes. Chi-square tests were used to test whether these variables were differently distributed between age groups, for males and females separately. Secondly, descriptive information in the form of numbers and percentages are presented on individual level sexual behaviour characteristics and data on all partnerships summarised at the individual level, for males and females separately. For categorical variables chi-square tests were used to test whether these characteristics were differently distributed between age groups for males and females separately. To analyse data on the age of sexual debut a survival approach was adopted with censoring at current age for those not yet sexually active. Log rank tests were used to compare the profiles of survival times between those of different age groups. Quantitative variables, including numbers of lifetime and recent partners were summarised with medians and inter-quartile ranges. Age distributions were compared using non-parametric tests for males and females separately.

The third part of the analysis focused only on individuals who reported at least one sexual partner during the past 12 months. Analysis was performed for males and females separately on a dataset consisting of all sexual partnerships reported by individuals of either sex pooled together, including up to a maximum of three partnerships for any one individual. For presentation purposes, partnerships were divided into those where sex had occurred more than five times during the previous year and those where sex had occurred five or less times. It was hypothesised that these partnerships may have different characteristics in terms of age difference between partners, resource exchange and condom use. Descriptive data are presented in the form of numbers and percentages for each partnership type and Chi-square tests were used to compare age distributions. It was further hypothesised that frequency of intercourse, age difference and resource exchange might be inter-related, and might each be related to condom use within partnerships.

Fourthly, the distribution of prevalent HIV by age among males and females, is presented graphically. A subset of the whole sample had data available on HIV serostatus and further analysis was restricted to this group. Logistic regression was used to assess the association between sociodemographic factors, HIV awareness and sexual behaviour characteristics and HIV infection among males and females in the study sample. Numbers and percentages of HIV positive individuals in each exposure class are presented. Logistic regression models were used to assess these associations. HIV was the outcome variable. A first set of models was run which were adjusted for age alone. Following this a model was run for each potential risk factor adjusted for age and village (for females) or village pair (for males, since there were no HIV infections in one village). Following this, a final model was constructed. This model included age, village (or village pair) and any other factors at the same or further up the hierarchical chain (to the left, as depicted in the framework diagram), that were retained in a stepwise model ran at each level. Thus, in this final model measures of association with HIV for sociodemographic factors were adjusted only for other sociodemographic factors, while sexual partnership characteristics could be adjusted for sociodemographic, HIV awareness and individual sexual behaviour characteristics that were retained in models run at each level.

4.05.4 The social epidemiology of sexual behaviour and HIV infection

The final results chapter, Chapter 8, presents measures of the association between the main exposure and outcome variables for the study. The principle measures of association presented in this thesis are odds ratios calculated from logistic regression and presented with 95% confidence intervals, although for one variable (ever had sex / age of sexual debut) hazard ratios are presented from Cox regression. It should be noted that when the occurrence of an outcome in the population of interest is rare, it has been shown that the odds ratio is numerically similar to the risk or prevalence ratio and can safely be interpreted in the same way. This is useful since prevalence ratios are more often referred to in common speech and are intuitively easier to understand (for example, "people with exposure A were twice as likely to be HIV positive than those without this exposure", is an example of use of a risk ratio). However, odds ratios are more often used in epidemiological studies, primarily because they can be calculated directly in case-control studies and because they are more easily calculated using the regression techniques widely used in modern epidemiology. In this study, odds ratios from logistic regression are presented for a wide variety of outcomes, but in many cases the outcomes under consideration cannot be considered 'rare'. The reader must therefore not interpret odds ratios in terms of prevalence or risk, but rather they must be interpreted directly as the ratio of the odds of having the outcome between exposure groups. Data on the number and proportion of individuals experiencing the outcome are presented in each chapter alongside odds ratios to aid interpretation.

Advantages of the use of logistic regression include that this technique can be relatively easily adapted to take account of the sampling scheme used and multiple confounders can be adjusted for in analyses. With regard to the sampling scheme, two issues were relevant to this study. Methods to take account of these issues were considered although ultimately a standard logistic regression strategy was employed. Firstly, all individuals within the sampling frame did not have an equal chance of being selected for inclusion in the study. This was because a fixed number of households were sampled per village (200) rather than using a sampling strategy that sampled households with probability proportional to village size. Village sizes ranged from 567 to 2256 households. Probability weighting might have been used to adjust the analyses for this. However, no weighting was applied to regression models in the analysis. It was felt unnecessary to apply such weighting since the villages themselves were purposively selected and not necessarily representative of the study region. Consequently, a weighted figure would not have much greater interpretative value than an unweighted one. Additionally, the primary aim of this thesis was not to provide population level estimates of indicators, but rather to examine the association between exposures and outcomes. It was not hypothesised that the relationships under study might differ by village. Regression models are presented with inclusion of a fixed effect term for village to account of variation in the distribution of outcome variables between villages. In cases where this was not possible since there were no outcome events in a given village, adjustment was instead made for village pair. These situations are clearly identified in the text and tables.

The second sampling issue was that standard logistic regression techniques assume that analytic units are independent of each other. This means that the outcome status of one sampling unit does not imply any information about the outcome status of any other sampling units. In this study, households were sampled, but individuals were included in the final analysis. Some households contained more than one individual included in the analysis. Additionally, in some parts of the analyses relating to sexual partnerships, individuals reported more than one partner, and all partnerships were pooled together in regression models. Since individuals from the same household may be more alike in their HIV risk profile, and sexual partnerships reported by the same individuals may be more alike than those reported by other individuals, outcome variables may have been clustered by either household or individual or both. In cases where responses are clustered by either household or individual the measure of effect calculated using standard techniques may be inappropriate and the 95% confidence intervals will have a tendency to be too narrow. Techniques are available for the computation of measures of effect and their confidence intervals that can account for this fact (212). However, these approaches were not used in this analysis. The rationale for this was two-fold. Firstly, such approaches add complexity to the analytic strategy, in particular because they are incompatible with the use of likelihood ratio tests to compare goodness of fit between models (in the case of generalised estimating equations with robust confidence intervals), or are computationally very complex and prone to providing invalid parameter estimates (in the case of random effects models). Secondly, it was felt that the likely impact of clustering in this study would be very small since the numbers of individuals of either sex sampled in each household cluster was quite low (only slightly over one on average), and the number of partnerships reported by individuals was also less than one on average. The sensitivity of the main findings of the study to these assumptions of independence are examined in Chapter 9 by presenting main findings of the study recomputed from logistic regression models that do take account of clustering.

An additional issue that needed to be considered when examining the precision estimates for measures of effect presented in this thesis was that of multiple testing. A large number of statistical tests were applied in the analysis. The application of multiple statistical tests generates an increased possibility that some relationships investigated will suggest statistical significance through chance. Techniques are available for the correction of estimates of statistical significance for the application of multiple tests. Such techniques generally tend to be conservative and have not been used in this study since it is intended as exploratory in nature. Nevertheless, care must be taken in interpreting the results of statistical tests, recognising the possibility for some relationships to be statistically significant simply due to sampling error.

A strategy was developed to guide the logistic regression analyses to be presented in Chapter 8 in line with the conceptual framework. The strategy was designed to appropriately take account of variables that might confound the associations of interest, to investigate potentially interesting interactions between variables in their association with outcomes, and to examine pathways of effect. Each of these three issues is discussed in more detail below.

Potential confounding variables are those that are associated with the exposure of interest, independently associated with the outcome of interest and do not lie on the causal pathway between exposure and outcome. As previously described key potential confounders were identified on which data was available. Potentially important confounding variables on which no data were available are identified in the discussion of this thesis. A large number of relationships between exposures and outcome were investigated. Logistic and Cox regression models were used to adjust estimates of effect for potentially confounding variables. In theory, each specific relationship studied could have been potentially confounded by a different range of factors. However, to simplify presentation and interpretation there needed to be some level of consistency between models in which factors were adjusted for. Since the inclusion of factors that do not in fact have a confounding influence on the relationship under study in regression models, as long as they are not on the relevant causal pathway, is most likely only to decrease the precision of presented estimates, an inclusive approach was taken. All models presented in Chapter 8 were adjusted for each of the confounding variables associated with any of the exposures. In practice, this meant that some factors were included in regression models, even though they may not have had a significant impact on the relevant odds ratio describing the association between the exposure and outcome under study.

The analytic approach adopted was to present data in such a way that the level of confounding can be interpreted by the reader. All relationships have been presented in at least the following manner. Count and percentage data are presented for all variables. Following this, an age-adjusted odds ratio has been presented by including a categorical term for age grouped into three-year age categories (14-16, 17-19, 20-22 and 23-25 years). Secondly, an odds ratio further adjusted for the confounding influence of other confounders is presented. The confounders included in the models were those that were hypothesised to confound the association between socioeconomic factors and HIV risk characteristics, and which were statistically significantly associated either with any exposure variable or with HIV infection in the vicariate analyses (as presented in chapters 8 and 9). This meant that adjustment was made for the number of young people co-resident in the household of the individual, whether or not they were still resident in the village where they were born and their orphanhood status. When considering socioeconomic role variables as exposures among females the odds ratio was also presented for all variables adjusted for the potential confounding influence of ever having had a child. Relevant measures of effect and 95% confidence intervals are presented both with and without adjustment for this factor, allowing the reader to judge the importance of reverse causality in determining the relationships observed. The association between socioeconomic roles and childbearing where this approach was not possible should be treated with particular caution. Finally, where the two socioeconomic role variables were the exposure of interest, an odds ratio was presented adjusted for the other socioeconomic role variable to assess the independence of these association with HIV risk outcomes.

In addition to consideration of confounding it was necessary to investigate the relationships for some interactions. Firstly, it was hypothesised that the relationships under study might have differed between younger and older individuals. Consequently, a logistic regression model for all exposure / outcome combinations was run with an interaction term for age divided into two groups (14-19 and 20-25). A likelihood ratio test compared the model with an interaction term with one without. Details of analyses stratified by age are provided in the tables for associations where $p<0.05$. A similar approach was adopted when investigating the presence of interaction between the socioeconomic variables in their association with HIV risk outcomes. Interaction between school attendance and temporary migrancy for each outcome was assessed with a likelihood ratio test comparing models (adjusted for confounders) both with and without an interaction term between these variables. Finally, interaction between each of the socioeconomic role variables and household wealth was also investigated in this manner. For this purpose, the two higher household wealth groups were grouped together to maximise the statistical power of the test.

The final statistical strategy issue to consider was that of the causal pathway between socioeconomic variables and HIV risk outcomes. In order to capture the full magnitude of the association between exposures and outcomes, the measures of effect between socioeconomic factors and HIV risk outcomes should not be “overadjusted” for variables that mark for the causal pathway between these. In this respect, analysis was guided by the conceptual framework, so that where characteristics of sexual behaviour were the outcome variables, estimates were not adjusted for HIV awareness variables pathway variables. Similarly, analyses with HIV as an outcome (in both Chapter 7 and 8) were not overadjusted for variables on the causal pathway.

A further research question however was the extent to which any association between socioeconomic factors and HIV risk outcomes was mediated by the hypothesised factors along the causal pathway. Analyses presented in chapter 8 explicitly sought to operationalise the framework, by seeking to identify whether adjustment for causal pathway variables altered the magnitude of the associations noted with socioeconomic factors (213). Regression models were first fitted without inclusion of factors on the hypothesised causal pathway. These pathway variables were then included to the model and the odds ratios from the two models were compared. Any change in the odds ratio noted between the socioeconomic factor and the HIV risk outcome upon inclusion of a pathway variable to the model would imply that some or all of the association between the exposure and outcome was mediated by differences in the distribution of the pathway variable between exposure groups. A particularly important question of interest in this regard was whether any association between household wealth and outcomes was mediated by the socioeconomic roles of young people.

Following these analytic principles, data are presented within Chapter 8 in three sections. Data are presented separately for males and females. Odds ratios in each table less than 1 represent lower risk characteristics among the "exposed" group for each comparison. "Exposed" categories were those from "poor, but a bit better off" and "doing OK" households (compared to the "very poor" group), those "attending school or college" (compared to those not doing so), and those who were "Temporary migrants" (compared to non-migrants). In each case, odds ratios are presented with 95% confidence intervals after adjustment for relevant confounders. Odds ratios and p-values from interaction tests that represent associations significant at $p<0.05$ are underlined, while those representing associations significant at $0.05<p<0.10$ are underlined with a dotted line. In each case, p-values were calculated from likelihood ratio tests comparing models with and without the factor of interest.

The tables present the distribution of key outcome variables by socioeconomic factors in five categories; HIV awareness variables, individual level sexual behaviour characteristics, partnership level sexual behaviour characteristics, co-factors (circumcision only) and the presence of HIV infection. Within each table, the population under study, and thus the denominator, is not always the same. Four separate denominator groups are used. The largest number of relationships have as the denominator group all individuals of each gender included in the final study sample. Two variables ("Did not want sex at debut" and "Continuing sexual activity") had as the denominator only individuals of the relevant ages who had previously been sexually active. Partnership level analyses were conducted on a denominator of all reported sexual partnerships during the previous year (as reported in tables 7.4 & 7.5). As such, odds ratios presented for these factors represent the association between exposures and these outcomes conditional upon individuals having been sexually active during the previous year. Finally, the data on HIV infection has as its denominator all individuals who had a serostatus result available, these being a subset of the full sample. Any additional missing data are detailed case by case in footnotes to the tables.

In the first section, the association between household wealth and HIV risk characteristics was investigated, using the methods described above to adjust for potential confounders and to investigate the interaction between wealth and age in these associations. Count and percentage data on the distribution of outcome variables are presented by household wealth in three categories developed from PWR alongside the age and confounder-adjusted odds ratios.

Secondly, the association between socioeconomic role variables and HIV risk outcomes is presented for each sex. Data are presented in the tables in the form of numbers, percentages and up to three adjusted odds ratios for males or four for females as described above. Significance test values from likelihood ratio tests for age interaction are also presented in these tables. This section also presents evidence on the interaction between the socioeconomic roles in their association with outcome variables. Finally, an assessment is made of whether HIV awareness or sexual behaviour variables act as pathway variables in explaining any of the associations noted. Variables that were on the causal pathway and that were themselves significantly associated with the outcome were included into models and the difference in odds ratio upon inclusion of these factors is noted.

The final section of the analysis goes on to examine the intersection between household wealth and individual socioeconomic roles. Firstly, it was hypothesised that the influence of individual socioeconomic role might vary according to background household wealth. A statistical test of significance from a likelihood ratio test comparing models with and without the interaction term is presented. Significance values from likelihood ratio tests are presented in a table and details of significant interactions discussed in the text.

The analysis the examined the extent to which any associations between household wealth and outcomes were mediated by differences in individual socioeconomic role variables. Odds ratios are initially presented for each of the two comparative wealth categories ("poor, but a bit better off" and "doing OK") adjusted for age and confounding variables. Following this, adjustment is made in turn for, firstly, attendance at school, secondly, the temporary migrancy variable and finally, both of these variables. The point estimate odds ratio in the new model is then presented. Confidence intervals are not presented for simplicity and since their magnitude will be very similar to those presented in the main model.

4.06 Statistical Power

As with all studies in which conclusions are based on statistical inference, the study is limited by the size of the sample in the precision of its estimates of the magnitude of association between exposures and outcomes. Table 5.1 presents calculations of statistical power to detect relationships of a given magnitude (expressed in terms of Odds Ratios) between each of the key hypothesised protective exposures and three key outcomes. Statistical power is defined as, [1 - the probability of not rejecting a null hypothesis on the basis of the sample taken when in fact the null hypothesis is in fact false], and is expressed as a percentage. The numbers of exposed and unexposed presented in the table are true values from the final study sample (see Chapter 5). Data are available on the outcomes among a different subgroup in each case, so N's vary widely for each comparison. Estimates of outcome frequency in the unexposed groups have been given realistic whole number values, based on the estimations of the prevalence of outcomes on the full study sample.

Table 4.6: Calculations of statistical power in examination of three major relationships under study in this thesis

Postulated protective exposure	Outcome	Population Subgroup	N (Exposed :Unexposed)	Estimated proportion of unexposed with outcome	Power to detect odds ratio		
					0.7	0.5	0.3
Males							
Household doing 'OK'	High number of lifetime partners (>3)	All sampled individuals	**120:796**	**35%**	**71.3%**	**97.8%**	**>99.9%**
In School	High number of lifetime partners (>3)	All sampled individuals	**628:288**	**35%**	**95.2%**	**>99.9%**	**>99.9%**
Non-migrant	High number of lifetime partners (>3)	All sampled individuals	**788:128**	**35%**	**77.5%**	**98.3%**	**>99.9%**
Household doing 'OK'	No condom use at last sex	All sexual partnerships *	**93:655**	**70%**	**30.9%**	**83.6%**	**>99.9%**
In School	No condom use at last sex	All sexual partnerships *	**403:305**	**70%**	**60.4%**	**99.4%**	**>99.9%**
Non-migrant	No condom use at last sex	All sexual partnerships *	**139:609**	**70%**	**38.2%**	**93.2%**	**>99.9%**
Household doing 'OK'	HIV positive	All individuals with oral fluid sample	**105:658**	**7%**	**5.7%**	**14.0%**	**32.9%**
In School	HIV positive	All individuals with oral fluid sample	**550:213**	**7%**	**16.2%**	**45.4%**	**79.8%**
Non-migrant	HIV positive	All individuals with oral fluid sample	**664:99**	**7%**	**11.1%**	**30.8%**	**60.8%**
Females							
Household doing 'OK'	High number of lifetime partners (>2)	All sampled individuals	**159:844**	**35%**	**83.0%**	**>99.9%**	**>99.9%**
In School	High number of lifetime partners (>2)	All sampled individuals	**623:380**	**35%**	**97.8%**	**>99.9%**	**>99.9%**
Non-migrant	High number of lifetime partners (>2)	All sampled individuals	**896:107**	**35%**	**71.7%**	**97.0%**	**>99.9%**
Household doing 'OK'	No condom use at last sex	All sexual partnerships *	**102:598**	**70%**	**32.8%**	**86.0%**	**>99.9%**
In School	No condom use at last sex	All sexual partnerships *	**346:354**	**70%**	**57.8%**	**99.1%**	**>99.9%**
Non-migrant	No condom use at last sex	All sexual partnerships *	**609:91**	**70%**	**26.5%**	**80.8%**	**>99.9%**
Household doing 'OK'	HIV positive	All individuals with oral fluid sample	144:736	**15%**	18.0%	54.7%	94.1%
In School	HIV positive	All individuals with oral fluid sample	555:325	**15%**	37.4%	85.6%	99.7%
Non-migrant	HIV positive	All individuals with oral fluid sample	798:82	**15%**	17.1%	48.1%	84.0%

* Up to a maximum of three reported partnerships during the last year for each individual. Shaded cells are those where the study generated <80% power to detect a statistically significant difference at p<0.05.

It can be seen from the table that the study generates relatively low power to detect associations of the order of Odds Ratio 0.7 for any of the outcomes. However, study power exceeds 80% to detect an odds ratio of 0.5 in outcome prevalence between exposed and unexposed groups in almost all cases. The exception to this is when HIV prevalence was the outcome. For this outcome, less than 80% power to detect an Odds Ratio of 0.5 is generated for all comparisons among males, and for estimations for the exposures household wealth and migrancy among females. Greater than 80% power was generated by the study to detect an Odds Ratio of less than 0.3 for all comparisons, except when HIV was the outcome among males.

Chapter 5. Description of sampled population

5.01 Introduction

Participatory wealth ranking was conducted between June and September 2001, while survey fieldwork was conducted between September and December 2001. Every effort was made in fieldwork to ensure a high response rate and thus to generate a study sample representative of the general population of the eight IMAGE study villages using the procedures described in the chapter above. However, as with all epidemiological studies there were inevitably some individuals who were eligible for inclusion in the study sample but who were not represented. Epidemiological studies need to assess the representatives of the data they collect in order to inform the interpretation of findings. This chapter describes the volume of data collected and the profile of individuals who were eligible for inclusion in the study sample and those who were included in the final study sample.

5.02 Objectives of chapter

1) Describe the data obtained by the study in terms of generating a sample frame and collecting wealth ranking information, enumerating selected households, and interviewing and collecting and analysing oral mucosal transudate samples from eligible individuals

2) Describe reasons why some eligible individuals were not included in the final study sample and assess any differences between those included in the study sample and those not included

3) Describe the socio-demographic characteristics of households and individuals included in the final study sample.

5.03 Results

5.03.1 Data obtained

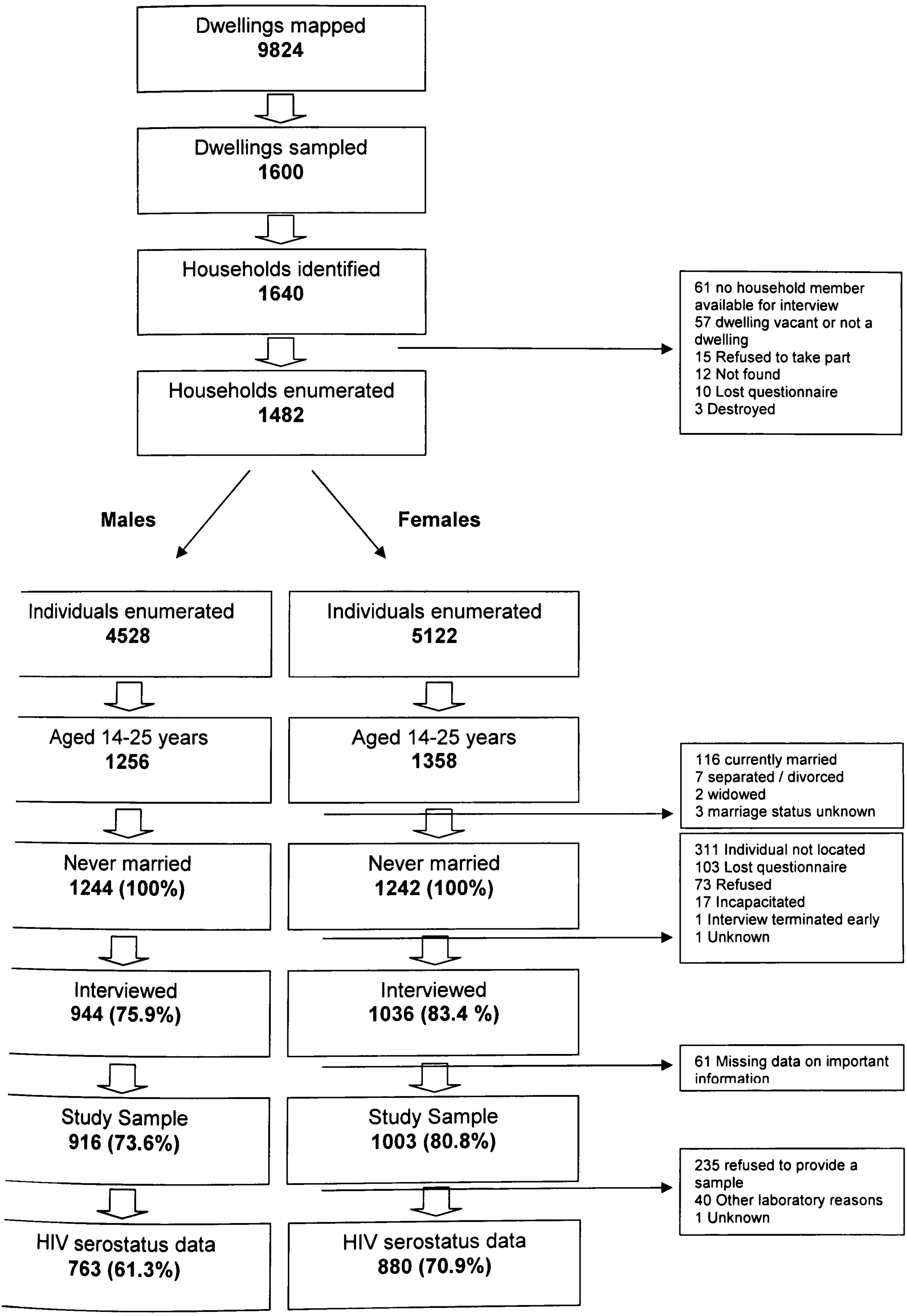
(a) Mapping and ranking

The wealth ranking and mapping process used to generate a sample frame identified 9824 dwellings in 79 village sections in the eight villages, and wealth ranking data was available for 9671 of these (98.4%). The number of dwellings identified in each village ranged from 567 to 2256.

Three ranking sessions occurred in each of the 79 village sections, thus giving 237 ranking sessions that occurred in total. Across all of the ranking sessions a total of 3556 *general statements* were coded describing the general properties of households seen as “very poor” (1240), “poor, but a bit better off” (1097) or “doing OK” (1216). A further 8856 *pile statements* were coded, describing the characteristics of the households included in each of the piles assembled by the wealth ranking process. In total, the *general* and *pile* statements were coded under 33 themes, and 880 statement codes. Some 81.1% of statements were assigned only one code. Of these, 319 *pile statement* codes were used on more than three occasions. Final *pile statement* scores were assessed for these statements and the index of household wealth was calculated as the average of these scores for all statements made about piles into which each household was ranked. Some 168 *general statement* codes were used more than three times. Bringing these together, 131 statement codes, under 23 themes, were made more than three times in both stages of the process. The full list of 131 statements is held in Appendix 3, and was used to identify the groupings of households identified in the next chapter. A condensed version, including only statements made more than 15 times in both phases, is held in Table 6.1 and discussed in Chapter 6.

A total of 200 dwellings were randomly sampled from each village irrespective of village size. Upon visit to 31 dwellings it was found that more than one household was resident at the dwelling. In 28 cases the sampled dwelling corresponded to two residential households, and there was one case each of three, four and eight households being identified at the mapped dwelling. Each of these individual households was eligible for inclusion in the study. The final sample of eligible households thus comprised 1640 households. Successful interviews were completed with 1482 (90.4%) of these households.

Figure 5.1: Data obtained in the study



A total of 9653 individuals were listed as members of the 1482 households for which a roster was available, of whom 53.0% were female (data on sex was missing for 3 individuals). The average household size was 6.5 people. Of this population, 3881 were aged 14-35 and eligible for inclusion in the main study, and of these approximately 2614 were aged 14-25. This is a close approximation since only year of birth was recorded on the household roster, so it was not possible to tell the precise age unless a young person's questionnaire was successfully completed. Of these individuals, 2486 had never been married. Three individuals were excluded because marriage status was unclear. Of the 125 individuals who had ever been married, only 11 were males, all of who were aged 20-25 years and were still currently married. Among females, 2 individuals were widowed, 7 had been separated and 105 were currently married. Among the females, 99 (86.8%) of those who had ever been married were aged over 20 years. Females who were married were much less likely to be attending school than unmarried individuals, even after adjustment for age. However, married females were no less likely to have completed secondary school (109/503, 21.7% among never marrieds vs. 18/99, 18.2% among those ever married) or to be temporary migrants (207/1133, 18.3% vs. 28/108, 25.9%). A simple age adjusted analysis also suggested no statistically significant relationship between ever marriage and having had more than 3 lifetime sexual partners (316/1035, 30.5% vs. 46/99 46.5%), condom use at last sex (193/460, 29.6% vs. 20/97, 20.6%) or being HIV positive (110/911, 12.1% vs. 13/78, 16.7%), with each of these associations attenuating after adjustment for age. However, ever married females all reported a single (spousal) partner during the past year, meaning that never married individuals were more likely both to report no or more than one sexual partner during the previous 12 months. Individuals who had ever been married were excluded from the remaining analyses.

A total of 1919 unmarried individuals (916 male, 1003 female) were successfully interviewed and comprised the study sample. The 916 males were from 652 households. In 451 cases only a single individual included in the final analysis was resident in a household. Additionally however, 148 households held two males included in the final sample, 44 households held three males, 8 held four and 1 had five males. The females were from 701 households, with 469 being the only female from a given household, 174 households having two included females, 48 households held three, 9 held four and 1 held six.

Finally, all 1919 individuals held in the final study sample were asked to provide a sample for HIV testing. HIV test results were available for 1643 individuals (85.6% of those interviewed, 66.1% of all those eligible).

5.03.2 Reasons for and factors associated with non-inclusion in the study sample

Successful interviews were completed with 1482 households, 90.4% of those eligible. The major reasons households were not included in the final sample were an inability to find a household member to conduct the roster interview with (n=61), or the dwelling was vacant or found not to be a household dwelling (n=57). Other reasons included refusal (n=15), the dwelling could not be found (n=12), the questionnaire was lost (n=10) or the dwelling was destroyed (n=3). Successful household interview rates ranged from 87.6% to

95.7% by village and there was little difference in the wealth ranking scores assigned to households that were and were not successfully interviewed. No other data on households that were eligible but not included in the study was available.

A total of 1919 unmarried individuals were successfully interviewed. The reasons for no interview being conducted were inability to interview the eligible individual after repeated household visits (n=311), the questionnaire was lost (n=103), refusal (n=73), the interviewee was incapacitated (n=17), the interview was interrupted (n=1) or the reason was unknown (n=1). A further 61 individuals were excluded from the analyses presented in this thesis as they had key questionnaire data relating to the main outcomes and exposures in this thesis missing. Data was missing on major sexual behaviour factors (n=24), HIV awareness (n=19), household wealth (n=13), the individual socioeconomic role (n=4) and key confounders (n=15). The final sample thus comprised 1919 individuals (c. 77% of those eligible) for whom data was available including 916 males and 1003 females. These individuals came from 1003 households.

A higher proportion of females than males were successfully included in the study sample (80.8% vs. 73.6%), while the proportion of eligible individuals successfully included in the final study sample varied from 67% to 84% between villages (Chi2 = 57.4, $p<0.001$). Tables 5.1 and 5.2 present key socio-demographic data on the 2614 eligible 14-25 year olds, 1919 interviewed individuals and 1643 individuals with HIV status data available, for males and females separately. Data on the reasons for non-inclusion in the study sample suggested that mobility was the most important reason for non-inclusion in the study sample. The socio-demographic profile also suggested that this was the case. The most striking difference between those with questionnaire data available and those eligible was the proportion sleeping away from home among both sexes (10.0% of those interviewed vs. 20.3% of those eligible for males, and 8.0% vs. 14.5% for females). Stated in a different manner, the study successfully interviewed 1747/2053 (85.1%) of those who were currently sleeping at home, compared with 172/433 (39.7%) of those who were sleeping away from home. This dynamic meant that those included in the final sample were also significantly younger, more likely to be students and less likely to have higher educational attainment among both sexes than those who were eligible but not included in the final sample. Additionally, interviewed individuals were less likely to have come from households with a larger number of 14-25 year olds.

Finally, HIV test results were available for 1643 individuals (763 males and 880 females, representing a total of 66.1% of those eligible). The major reason for no HIV test result being available was refusal to provide an oral fluid sample (n=235/276, 85.1% of all those individuals for whom no result was available). A smaller number of samples were unprocessable, lost in transit, or other data errors prevented their inclusion (n=40), while in one case the reason was unknown. A second issue was also noted in the conduct of the study. A subset of samples were not processed within 21 days of collection as recommended by the manufacturer (n=598, 36.4%). Of these samples, 301 were processed within 25 days of collection (50.3%), and a further 212 within 30 days of collection (35.4%), with only a small subset apparently processed outside this period (n=85, 14.2%). Data on the validity of the results of samples processed outside 21 days is not available. This issue is addressed further in Chapter 9. Results for all samples were included in the main analysis.

Individuals who were interviewed and also had an HIV serostatus test result available also had a slightly different profile when compared to those who only consented to the interview. Individuals with an HIV test result were also more likely to have been sleeping away from the home among both sexes (9.9% vs. 9.0% for the males, 8.0% vs. 7.0% for females). Among males, additionally those who had travelled to a city were less likely to have an HIV test result available, while individuals with an HIV test result were also more likely to be from the young groups, to be students and to have lower levels of education (see Table 5.1). Among females, those whose first language was not Sepedi, those who had not travelled to a city in the previous year and those who had been less mobile were each more likely to have a biological sample available for analysis.

Table 5.2 Socio-demographic characteristics of males who were eligible for the study, interviewed and had data on HIV status available

		Eligible (N=1244)		Questionnaire data (N=916)		P[1]	Questionnaire and lab data (N=763)		P[2]
		n	%	n	%		n	%	
Age group	14-19	691	55.6	577	63.0		504	66.1	
	20-25	553	44.5	339	37.0	<0.001	259	33.9	<0.001
Sleeping away from the home	No	991	79.7	824	90.0		693	90.8	
	Yes	253	20.3	92	10.0	<0.001	70	9.2	0.051
Main current activity	Self employed	16	1.3	13	1.4		9	1.2	
	In regular work	92	7.4	49	5.4		34	4.5	
	Irregular work	53	4.3	35	3.8		28	3.7	
	Student	785	63.2	628	68.6		550	72.1	
	Unemployed	272	21.9	178	19.4		133	17.4	
	Not active	25	2.0	13	1.4	<0.001	9	1.2	<0.001
Educational attainment	None	8	0.6	3	0.3		3	0.4	
	Attended primary	301	24.2	228	24.9		193	25.3	
	Attended secondary	786	63.2	594	64.9		501	65.8	
	Completed second.	98	7.9	64	7.0		49	6.4	
	Attended higher	50	4.0	26	2.8	<0.001	16	2.1	0.015
Number of co-resident 14-25 year olds	1 or 2	488	39.2	381	41.6		325	42.6	
	3 or more	756	60.8	535	58.4	0.005	438	57.4	0.170
Language	Tsonga	-	-	38	4.2		33	4.3	
	Sepedi	-	-	232	90.8		694	91.0	
	Other	-	-	46	5.0	-	36	4.7	0.552
Months staying in the village in past year	More than 6mo.	-	-	837	91.5		701	91.9	
	6mo. Or less	-	-	78	8.5	-	62	8.1	0.333
Travel to a city in past year	No	-	-	617	67.4		529	69.3	
	Yes	-	-	298	32.6	-	234	30.7	0.006
Born in village of current residence	Yes	-	-	701	76.5		581	76.2	
	No	-	-	215	23.5	-	182	23.9	0.543
Orphan Status	Both parents alive	-	-	598	65.3		507	66.5	
	One parent dead	-	-	290	31.7		233	30.5	
	Both parents dead	-	-	28	3.1	-	23	3.0	0.248
Position in household	Head	-	-	4	0.4		2	0.3	
	Offspring of head	-	-	701	76.5		581	76.2	
	Other blood relative	-	-	196	21.4		167	21.9	
	Unrelated / related only by marriage	-	-	15	1.6	-	13	1.7	0.271
Ever had sex	No	-	-	338	36.9		286	37.5	
	Yes	-	-	578	63.1	-	477	62.5	0.413
Non-spousal partner during the past year	No	-	-	393	42.9		329	43.1	
	Yes	-	-	523	57.1	-	434	56.9	0.769

[1] Chi-square test comparing those interviewed with those eligible but not interviewed

[2] Chi-square test comparing those for whom HIV serostatus data are available with those who were interviewed but no HIV data are available

Data was missing for the following variables; main current activity (1), educational attainment (1), months in the village (1), overnight trip to a city (1).

Table 5.3 Socio-demographic characteristics of females who were eligible for the study, interviewed and had data on HIV status available

		Eligible (N=1242)		Questionnaire data (N=1003)		P[1]	Questionnaire and lab data (N=880)		P[2]
		N	%	n	%		n	%	
Age group	14-19	739	59.5	621	61.9		557	63.3	
	20-25	503	40.5	382	38.1	<0.001	323	36.7	0.016
Sleeping away from the home	No	1062	85.5	923	92.0		819	93.1	
	Yes	180	14.5	80	8.0	<0.001	61	6.9	0.001
Main current activity	Self employed	21	1.7	15	1.5		11	1.3	
	In regular work	40	3.2	25	2.5		22	2.5	
	Irregular work	32	2.6	23	2.3		20	2.3	
	Student	758	61.1	623	62.1		555	63.1	
	Unemployed	374	30.1	306	30.5		265	30.1	
	Not active	16	1.3	11	1.1	0.015	7	0.8	0.068
Educational attainment	None	7	0.6	5	0.5		3	0.3	
	Attended primary	224	18.0	187	18.6		165	18.8	
	Attended secondary	881	70.9	717	71.5		637	72.4	
	Completed second.	89	7.2	68	6.8		59	6.7	
	Attended higher	41	3.3	26	2.6	0.030	16	1.8	<0.001
Number of co-resident 14-25 year olds	1 or 2	471	37.9	398	39.7		348	39.6	
	3 or more	771	62.1	605	0.3	<0.001	532	60.4	0.815
Language	Sepedi			913	91.0		795	90.3	
	Tsonga			39	3.9		35	4.0	
	Other			51	5.1		50	5.7	0.062
Months staying in the village in past year	More than 6mo.			942	94.0		833	94.8	
	6mo. Or less			60	6.0		46	5.2	0.007
Travel to a city in past year	No			743	74.1		663	75.3	
	Yes			260	25.9		217	24.7	0.015
Born in village of current residence	Yes			740	73.8		641	72.8	
	No			263	26.2		239	27.2	0.071
Orphan Status	Both parents alive			681	67.9		608	69.1	
	One parent dead			297	29.6		251	28.5	
	Both parents dead			25	2.5		21	2.4	0.095
Position in household	Head			5	0.9		4	0.5	
	Offspring of head			776	77.5		685	77.9	
	Other blood relative			200	20.0		171	19.5	
	Unrelated / related only by marriage			21	2.1		19	2.2	0.668
Ever had sex	No			297	29.6		263	29.9	
	Yes			706	70.4		617	70.1	0.610
Ever had a child	No			641	63.9		562	63.9	
	Yes			362	36.1		318	36.1	0.937
Non-spousal partner during the past year	No			367	36.6		325	36.9	
	Yes			736	63.4		555	63.1	0.548

[1] Chi-square test comparing those interviewed with those eligible but not interviewed

[2] Chi-square test comparing those for whom HIV serostatus data are available with those who were interviewed but no HIV data are available

Data was missing for the following variables; main current activity (1), educational attainment (1), months in the village (1), position in the household (1).

5.03.3 Socio-demographic profile of individuals in the final study sample

Table 5.3 presents data on the structure of the households that individuals aged 14-25 years came from. Households containing 14-25 year olds were larger than the whole sample (mean household size = 7.4 persons). The heads of the household were generally adults of working age, with 262 (26.2%) household heads over 60 years. The household head was most often male, although a significant proportion of households were headed by a female (n=402, 40.1%). Female household heads were more likely to be over 60 years than their male counterparts (134, 34.2% of female households head vs. 122, 20.7% of male household heads), and were also more likely to be currently sleeping at the home (343, 87.1% vs. 368, 62.4%).

A high proportion of household heads were not currently sleeping in the home (n=278, 27.7%). Households usually contained young people living alongside older adults and / or pensioners, with a very small number of households comprising only young people with or without children (n=4). Households were generally structured on a 'nuclear' basis (parents with children), although a wide number of variants of this existed, including couples living alone, unmarried parents with children, or nuclear homes with extended family members. However, other types of household were common, for example unmarried-headed, extended family households were common (n=329, 32.8%).

Table 5.4 : Characteristics of households included in the final study sample

		Households (N=1003)	
		N	%
Household Size	1 or 2	21	2.1
	3-5	243	24.2
	6-10	594	59.2
	11+	145	14.5
Age of head	18-25	10	1.0
	26-45	275	27.5
	46-60	453	45.3
	61+	262	26.2
Sex of head	Female	402	40.1
	Male	601	59.9
Migrancy Status of head	Not currently sleeping away from home	725	72.3
	Currently sleeping away from home	278	27.7
Age structure	Young people with or without kids, no adults or pensioners	4	0.4
	Adults with children and/or young people	671	66.9
	Pensioners with children and/or young people	48	4.8
	Adults and pensioners with children and/or young people	280	27.9
Household Structure	Single person	1	0.1
	Couple	0	0.0
	Single parent nuclear	99	9.9
	Nuclear	225	22.4
	Extended nuclear	333	33.2
	Non-nuclear partnerships with extended households	16	1.6
	Single-headed, non-nuclear extended households	329	32.8
Households with at least one member	Receiving a pension	399	40.1
	Student	947	95.0
	Not working, but willing and able to work	737	73.9
	In irregular work	183	18.4
	In self employment, domestic work or agriculture	366	36.7
	In regular or salaried employment (sleeping at home)	300	30.1
	In regular or salaried employment (not sleeping at home)	387	38.8
Household Income Portfolio	No self-generated sources of income	37	3.7
	Irregular incomes only	20	2.0
	Self employment / domestic work with or without irregular incomes	123	12.4
	Pensions only	94	9.5
	Pensions w. self employment and/or irregular incomes	107	10.8
	Non-resident salaried employment only	127	12.8
	Non-resident salaried employment w. self employment, pensions and / or irregular incomes	186	18.8
	Resident salaried employment only	102	10.3
	Resident salaried employment w. any other incomes	194	19.6
Most Important Income Source	Financial help from outside the household	64	6.6
	Migrant Labour	216	22.4
	Pensions or grants	265	27.5
	Local Labour	178	18.5
	Irregular incomes	30	3.1
	Self employment	211	21.9

Data was missing for the following variables; age of household head (3), households with at least one member receiving any income (6) or receiving a pension (7), income portfolio (13), most important income (39).

Most households contained at least one student (95.0%) and an individual who was not a student, able and willing to work but unemployed (73.9%). Some 40.1% of households had at least one member receiving a grant or pension. More households contained a migrant worker than a locally employed individual, with 38.8% of households having a member working in regular paid employment, but not currently sleeping in the home, while 30.1% of households held a regularly working individual currently staying in the homestead. Self-employment, usually in the informal sector, was a common household livelihood tactic, with 36.7% of households having at least one member engaged in this, most often adult women.

When households were classified on the basis of their combined portfolio of individual income sources, just under 4% of households reported no members receiving any income from work or grants / pensions. A further 2.0% reported only incomes generated from irregular work. Self employment / domestic work in the absence of either grants / pensions or regular employment supported a further 12.4% of households. Some 9.5% of households were dependent on grants / pensions alone, while another 10.8% combined these with incomes from either self employment / domestic work or casual labour. 31.6% of households held members who had regular employment but were not currently at home, and this was either in isolation or combined with other income sources. Finally, 29.1% of homes had access to local regular employment, either alone (10.3%) or combined with other income sources (18.8%).

Grants / pensions were the most important source of income for over one quarter of households. State old age pensions were by far the most important form of state grant, although private pensions and child grants also played a role (data not shown). Migrant labour was a more common most important source of income than locally earned incomes. Employment in mining was the commonest sector of employment for those staying both at home and away from the home. Self employment activities such as hawking, running a *shebeen* (alcohol bar) and sewing / mending clothes were important income generating activities among those who were self employed. Of those receiving help from outside the household, this most often came from parents.

Of individuals with questionnaire data available, slightly more 14-19 year olds were included in the study than 20-25 year olds for both sexes. For males, the relative proportion of 14-19 year olds in the final study sample was slightly higher than for females (62.9% vs. 62.2%). The majority of those interviewed were still in school (68.6% of males and 62.1% of females). Almost all individuals had attended primary school, the majority had attended secondary school (particularly among the older age group, data not shown), but relatively few had completed the secondary school qualification (9.8% of males and 9.4% of females). Mobility was relatively common. Among males, nearly a quarter of individuals had been born in a place outside the current village of residence, while nearly 1/3 had made a trip to a city and some 8.5% had spent 6 months or less sleeping at the home in the past year. A slightly higher proportion of females had been born outside the village of current residence (26.2% vs. 23.5% of males). However, a slightly smaller proportion of females had been mobile in the past 12 months by all measures. Single parent orphanhood was relatively common among the study sample, with 31.7% of males and 29.6% of females reporting this. The absent parent was more often the father than the mother for both sexes, with 256/290 (90.1%) of male single parent orphans having lost a father, while for females the figure was similar (261/297, 91.3%). Most individuals were either the offspring, or another blood relative of the current household head. Most had been sexually active, and a majority had had a non-spousal sexual partner during the course of the previous 12 months. A significant number of females had previously had at least one child (36.1%).

5.04 Summary of key findings

Study procedures generated a sampling frame of 9824 households in 8 villages of which 1640 were sampled and 1482 were enumerated. Some 2614 individuals (1256 male and 1358 female) aged 14-25 were enumerated of whom a small proportion were married, of whom more were female than male. Some 916 (73.7%) eligible males and 1003 (80.9%) eligible females were successfully interviewed, of whom 1646 (85.6%) provided an oral fluid specimen that was successfully tested for antibodies to HIV. Those who were not interviewed were more likely to be sleeping away from the home, which in turn meant these individuals were more likely to be older, male and out of school. The study sample contained individuals who were likely to be in school or unemployed, had relatively high levels of mobility, over half were sexually active and a significant proportion of females had had a child.

Chapter 6. Socioeconomic factors

6.01 Introduction

After the study data had been collected, entered into databases and appropriately cleaned the next stage of the analysis was to gain greater insight into the distribution of key socioeconomic factors within the study sample. As described earlier, a key aim of the study was to generate a robust measure of household wealth using the available data from PWR and / or the survey procedures, and to describe how individual socioeconomic roles were distributed within the population. This chapter describes the results of these investigations.

6.02 Objectives of chapter

1) Describe indicators and local perceptions of relative wealth for the study population
2) Develop an indicator of household wealth from three different techniques and assess their agreement before identifying one as a measure of household wealth for use throughout the remainder of the study
3) Describe the distribution of individual socioeconomic roles within the study sample, in particular attendance at school or college and temporary migration
4) Assess the association between household wealth and socioeconomic roles among individuals of either sex.

6.03 Results

6.03.1 Distribution and determinants of wealth

Table 6.1 shows the 47 statements made more than 15 times during both the *general statements* and the *pile statement* phase of the wealth ranking. In the left hand two columns of the table are given the number of times each statement was made in total during the *pile statement* phase of the ranking procedure and the associated average *pile statement* score. The statements in the table are listed in ascending order of their *pile statement* scores, so that those statements made most often made about piles of the poorest households appear at the top of the table. On the right of the table is shown the frequency with which each statement was made in answer to questions about the general properties of households that were "doing Ok", "poor, but a bit better off", and "very poor". By ordering statements relating to local perceptions of economic well-being made during PWR by their *pile statement score*, table 6.1 provides an overview of the range and construction of welfare among households in the study sample.

Table 6.1: Pile statement scores and frequency of statements made during participatory wealth ranking in rural South Africa, in descending order of pile statement score

Pile statements		Theme	Statement	General statements: No. of times said		
Pile statement score	No. of times said			Very poor	Poor but a bit better off	Doing OK
0.0	22	Family and household	Orphanhood/no parents	24		
0.4	39	Food	Beg for food	33		
1.1	85	Begging	Begging	49		
3.1	134	Food	No food	137		
3.2	41	Housing	Not got shelter	33		
3.7	58	Employment	No one is working	34		
5.2	101	Schooling	Doesn't go to school	39		
5.6	73	Clothing	No clothes/do not have clothes	73		
5.7	199	Employment	Not got job(s)/unemployed	113		
5.8	22	Food	Sleep without food	17		
6.4	82	Money	Don't have/earn money/income	49		
8.3	100	Schooling	Unable to/can't afford to go to school	66		
9.0	67	Housing	Not got housing	65		
9.6	37	Schooling	Cannot afford/does not pay school fee	18		
11.9	23	Clothing	Tattered/torn/poor clothes	20		
14.4	76	Housing	Shacks	18		
15.0	51	Housing	No proper housing/shelter	18		
22.7	64	Housing	Bad/poor housing	19		
24.0	175	Employment	Farms		80	
28.4	145	Self employment	Selling fruits and vegetables		39	
28.5	71	Food	Mealy meal only		37	
28.7	99	Employment	Domestic work		45	
29.9	60	Pensions	Pension and many responsibilities		25	
34.7	26	Food	At least have food		19	
35.1	28	Food	Little food		33	
35.7	47	Self employment	Self employed		17	
38.2	55	Clothing	Second-hand clothes		21	
39.4	70	Money	Little money/income/earn less		29	
39.8	44	Housing	Small/little housing		26	
40.4	25	Schooling	Attains Matric/ std 10/grade 12		17	
44.4	70	Pensions	Receiving pension		16	
61.4	79	Schooling	Able to/ affords to go to school		29	
65.4	28	Employment	Got jobs/employed		24	18
71.0	32	Clothing	Good clothes			58
78.6	26	Clothing	Children have good clothes			30
80.8	134	Self employment	Taxis			41
83.1	104	Cars	Have/drive cars			50
84.7	101	Employment	Government			26
84.8	162	Schooling	Attains university/tertiary			52
86.4	97	Employment	Both husband and wife employed			18
87.9	163	Housing	Big house			96
90.1	123	Schooling	Private/expensive			76
90.4	73	Housing	Beautiful/attractive housing			42
93.8	65	Self employment	Has a business			47
95.5	74	Self employment	Shop owners			32
95.6	142	Cars	Have/ drive expensive/flashy cars			102
95.7	47	Housing	Tiled housing			21

The ten commonest themes raised during PWR discussions (see Table 4.3) are each represented in this table. Additionally, the theme "Begging" was represented in this table, but was the eleventh commonest theme overall. The commonest statements included those relating to both the determinants (employment, self employment, money, family and household, pensions) and the outcomes of increased or decreased wealth (food, schooling, clothing, housing, cars, begging). Statements at the top of the table, associated with pile statement scores of approximately 0-24 were most commonly made in relation to households deemed "very poor" by PWR participants. These statements suggested households that were struggling to survive. Such households were struggling to feed themselves and to clothe or educate their children. They had little or no access to jobs or housing. Further down the table were statements associated with *pile statement* scores of approximately 24 to 70, and which were made most commonly in relation to properties of households deemed "poor, but a bit better off" by study participants. Such households had access to relatively low paid jobs, such as domestic work, temporary employment and hawking or selling food at the side of the road. They exhibited a basic ability to meet basic human rights (food, education). Finally statements associated with households that were "doing OK" (scoring between approximately 70 and 100) suggested that some households were much better equipped for survival among the study population. These households had access to good food, drove cars and had big or attractive housing. Some members were employed in high–return and/or high-stability activities such as mine work, teaching, government employment or owning taxis or local shops.

Table 6.2 shows the distribution of each of the indicators of household wealth used in the construction of the two indices of wealth that used survey data, among the 1003 households that individuals in the final study sample came from. Modern assets were widely distributed within the population. However, 26.6% of households reported owning none of the listed assets, while those in the top quintile owned assets with a value in excess of 1100 ZAR per person. Livestock assets were also widely distributed. The majority of households were built of relatively simple materials. Some 17.2% of households had no access to a toilet, with very few households having a modern flush toilet. Electricity supply was determined largely by village of residence, with two villages remaining largely unelectrified. Conversely, accessibility of water was problematic for the majority of the population, with only a very small number of households having access to a yard pipe.

Table 6.2: Distribution of indicators of household wealth from survey data

Indicator	Group	N	%
N		1003	
Estimated value of selected non-livestock assets per person [a]	0 ZAR	264	26.6
	1-132 ZAR	142	14.3
	132-350 ZAR	225	22.6
	350-1100 ZAR	171	17.2
	> 1100 ZAR	192	19.3
Estimated value of selected livestock assets per person [a]	0 ZAR	261	29.0
	1-220 ZAR	92	9.2
	220-1115 ZAR	242	24.2
	1115-2440 ZAR	219	21.9
	> 2440 ZAR	158	15.8
Land tenure	No	718	71.7
	Yes	283	28.3
Quality of house wall material [a]	Poor	542	54.0
	Good	461	46.0
Quality of toilet facility	No facility	172	17.2
	Basic	818	81.6
	Modern	12	1.2
Household Electricity	No	308	30.8
	Yes	693	69.2
Accessibility of water supply	Low	343	34.3
	Medium	616	61.5
	Good	42	4.2
Density of household living conditions[a]	>1rm per person	833	83.2
	<=1rm per person	168	16.8
Proportion of household members receiving a regular income [a]	0	181	18.1
	Less than 25%	465	46.4
	25-49%	285	28.4
	50% or more	72	7.2
Educational level of household head	No schooling	390	38.9
	Attended primary	397	39.6
	Attended second	216	21.5
Percentage of household members working age adults	50% or less	291	29.2
	>50%	706	70.8
Gender of household head	Female	402	40.1
	Male	601	59.9
Second most important household income	Non-Financial	365	36.4
	Financial	638	63.6
Regularity of household having a meal consisting of mielie meal alone, bread alone or worse	Often	380	37.9
	A few times	287	28.6
	Once only	87	8.7
	Never	249	24.8
Car ownership	No	801	79.9
	Yes	201	20.1
Schooling (7-13yrs)	Any not attending	27	2.7
	All attending	706	97.3
Schooling (14-19yrs)	Any not attending	163	18.9
	All attending	698	81.1
Schooling (20-25yrs)	All not achieved college or techikon	601	94.5
	Any achieved college or technikon	62	5.5

[a] denotes variables that were grouped for presentation in table, but where an ordered or continuous variable was used in the PCA analysis.
ZAR=South African Rand

Four variables relating to household income or human capital were considered. Some 18.1% of households had no adults receiving an income from regular work, self employment or a pension, while many households were headed by an individual with no education (38.9%). The food security variable suggested high levels of food insecurity in the study sample with 37.9% of households reporting that they often had a meal consisting only of the most basic staple ingredients. Cars were owned by 20.1% of households. School attendance was high for young children but lower at later ages.

6.03.2 Indices of household wealth

In the PWR process between 4 and 11 ranking piles were formed in each session, with 6 being the most common. In most ranking sessions the number of households ranked in the poorest piles was higher than those in the wealthiest piles. For the PWR method it was also possible to take two further analytic steps, as described in the methods. Firstly, a visual assessment was used to group the households into three wealth groups. From table 6.1, it can be seen that relatively clear groups of scores were related to the three different levels of wealth as described by PWR participants ("very poor", "poor, but a bit better off", "doing OK"). When all statements made over three times were included in a similar table, cut-off values between these groups were identified at scores of 23.9 and 70.2 (see Appendix 3). Mapping these scores onto the household wealth index from PWR suggested a distribution of 309 households (30.8%) classified as "very poor" (scoring 0-23.9), 564 households (56.2%) classified as "poor, but a bit better off" (scoring 23.9-73.5) and 130 households (13.0%) classified as "doing OK" (scoring 73.5-100).

The second additional step that was taken with the PWR method was that the agreement between the separate scores from three rankings was estimated. An identical score in each of the three ranking processes may have slightly different interpretation, depending on both the number of ranking piles and the descriptions of poverty assigned to each pile. Nevertheless, an estimation of the intraclass correlation coefficient for these three scores gives an estimate of the consistency with which the scores agree with each other. The single measure intraclass correlation coefficient from a random effects two-way ANOVA was 0.81 (95%CI 0.79-0.82), denoting a high degree of agreement.

For Method 2, data from the PWR process and survey were linked. Broadly, where relevant PWR statements scored less than 24, the most related survey item was given a score of -2, and where relevant statements scored more than 70 the associated survey item was scored 2. A sliding scale for intermediate situations was developed for each item where this was possible, though the decision as to how to develop the scale was inevitably subjective. For school attendance, data was stratified on the basis of age since this was strongly associated with school attendance and the level of schooling attained. The scoring system and its relation to the PWR statements is shown in Table 6.3.

Table 6.3: Statement scores for poverty statements from PWR and scores for indicators collected in the survey data to create household index of wealth (Method 2)

Theme(s)	Relevant statements (score)	Relevant data from survey	Score applied to survey data
Employment, Self employment, Pensions, Money	Shop owners (95.5) Has a business (93.8) Both husband and wife employed (86.4) Government (84.7) Taxis (80.8) Got job / employed (65.4) Receiving pension (44.4) Self employed (35.7) Pension and many responsibilities (29.9) Domestic work (28.7) Selling fruits and vegetables (28.4) Farms (24.0) Don't have/earn money/income (6.4) Not got job(s)/unemployed (5.7) No one is working (3.7)	More than one household member has a salaried job *Either* one household member has a salaried job, *or* three or more have a pension or other work No household members have a salaried job, but two has a pension or other work No household members have a salaried job, but one has a pension or other work No household members have a salaried job, pension or other work	2 1 0 -1 -2
Schooling	Private/expensive (90.1) Attains university/tertiary (84.8) Able to/ affords to go to school (61.4) Attains matric (40.4) Can not afford / doesn't pay school fees (9.6) Unable to/can't afford to go to school (8.3) Doesn't go to school (5.2)	If there are 20-25 year olds, if any attending or already achieved technikon or university If there are 14-19 year olds and all are in school If there are 7-13 year olds and all are in school **OR** If there are 14-19 year olds and any are not attending school **OR** If no 7-25 year olds in household If there are 7-13 year olds and any are not attending school **Overall score; if there were young people from more than one age group in the household the average of the three scores was used**	2 1 0 -2
Housing	Tiled housing (95.7) Beautiful/attractive housing (90.4) Big house (87.9) Small / little housing (39.8) Bad/poor housing (22.7) No proper housing/shelter (15.0) Shacks (14.4) Mud housing (13.3) Not got housing (9.0) Not got shelter (3.2)	Face bricks Block bricks with cement covering Mud bricks, or block bricks without cement covering Tin or mud and sticks	2 1 0 -2
Food, begging	Little food (35.1) At least have food (34.7) Mealy meal only (28.5) Sleep without food (5.8) No food (3.1) Begging (1.1) Beg for food (0.4)	Food insecurity score 2-3 Food insecurity score 4 Food insecurity score 5-6 Food insecurity score 7-8 **Sum of two questions about the frequency of poor food security during the last month*** pre-scored as Never (1), Once only (2), A few times (3), Often (4).	1 0 -1 -2
Cars	Have/ drive expensive/flashy cars (95.6) Have/drive cars (83.1)	Own any cars	2
Family and Household	Widows 1.8, n=15^ Orphanhood/no parents (0.0)	Female Headed Household AND / OR Household consists only of children / young people	-2

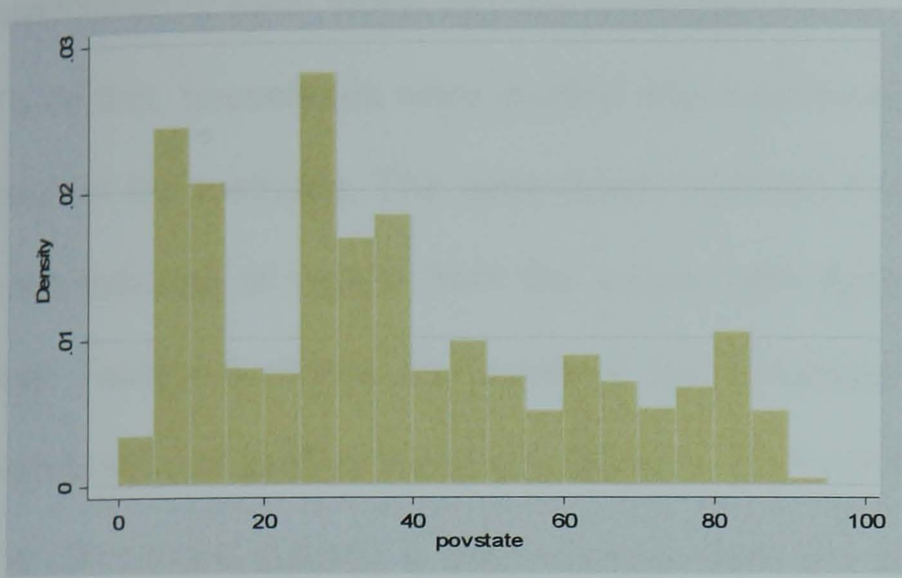
* The two questions were During the last month how often a) have most of the family had a meal that consisted of pap alone, bread alone or worse , and b) have <u>you or any of your own children</u> gone without food or had a reduced amount to eat for a single day because of a shortage of food ?

^ This statements is not listed in Table 6.1 because it was made less than 15 times in one stage, but was the second most common single statement about family and household made during the PWR process

During the assembly of the principal components model (Method 3), three factors (percentage of household members of working adult age, quality of water supply and land tenure) were excluded because they were not significantly associated with the value of non-livestock assets. A further two factors were excluded during the assembly of the PCA model since they had low component loading factors (presence of a household electricity supply, and the nature of the second most important ranked household income). In the final principal components analysis, the first principal component explained 22.7% of the variance of the factors in the model. The next two components explained 16.6% and 12.9% of the variance respectively. The greatest weight in the first principal component was given to the density of household living conditions (scoring coefficient = 0.44), with the value of non-livestock assets (0.42) and the food security indicator (0.39) also being important. The lowest component loading factor was given to the proportion of individuals receiving an income (0.23).

The frequency distribution of the three indices of household wealth generated among the 1003 households included in the final sample is shown in Figure 6.4. The available data made it possible to generate a welfare index for all 1003 households from the PWR data, for 989 households for the combined method (Method 2) and 991 households from the PCA method (Method 3). It can be seen that both methods 2 and 3 produced relatively normal distributions, while the PWR method was skewed slightly to the right suggesting a higher bulk of households at lower levels of wealth.

Figure 6.4: Frequency distribution of indicators of household wealth



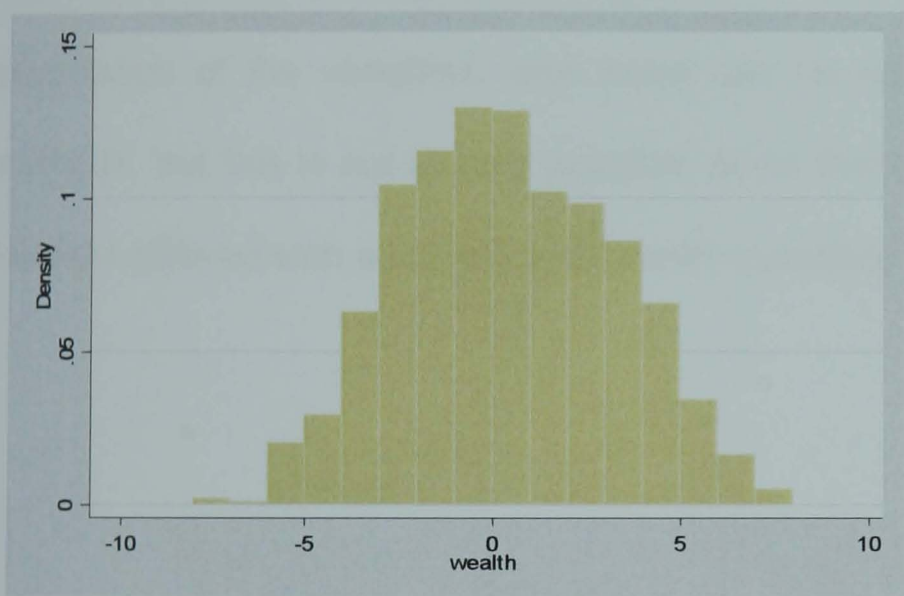
PWR only (Method 1)

N: 1003

Mean: 37.4

SD: 24.0

IQ Range: 16.2-53.4



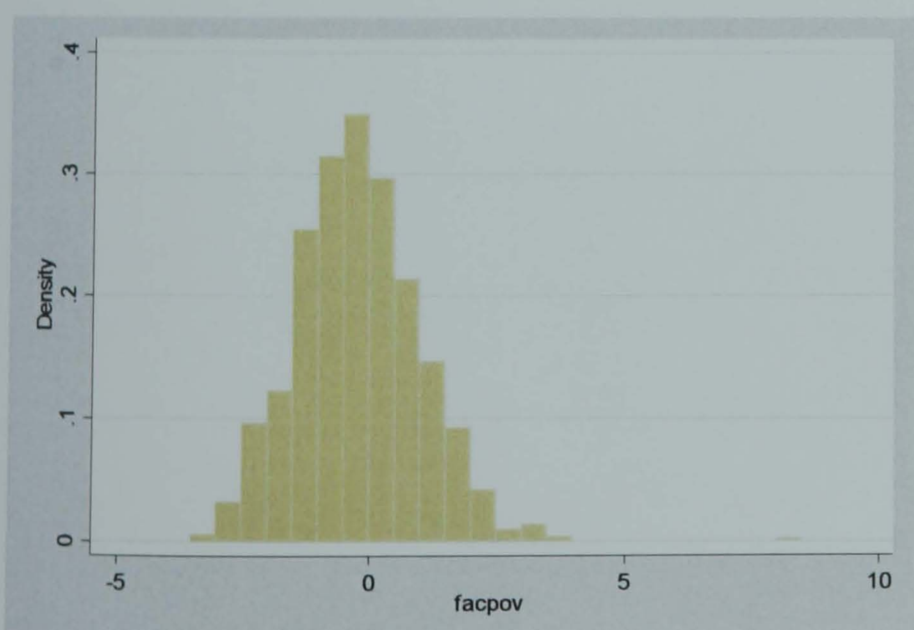
PWR and survey (Method 2)

N: 989

Mean : 0.2

SD : 2.9

IQ Range : -2 – 2.3



PCA (Method 3)

N:991

Mean: -0.2

SD: 1.2

IQ Range:-1.0 – 0.6

After generating and examining the distribution of the three indices, the association of each index with the individual survey indicators was estimated. To do this, households were divided into wealth-rank terciles on the basis of each of the methods. The association between these terciles of wealth and each indicator of wealth from the survey was assessed using a Chi-square test. Table 6.5 shows the value of the Chi-square statistic and statistical significance of each of these associations. It should be noted that the value of the Chi-square statistic is dependent on both the strength of the association and the number of degrees of freedom in the contingency table of the association of the variables, and these can be compared across the three methods, but this is not always possible down the table since the degrees of freedom differed with each index of wealth included.

Table 6.5: The association between household wealth rank (in tercile groups) and indicators of socioeconomic status collected in the survey

	Method 1 : PWR		Method 2 :Survey + PWR		Method 3 : Survey only	
	Chi-square	P	Chi-square	P	Chi-square	P
Estimated value of selected non-livestock assets per person	91.9	<0.001	328.1	<0.001	341.4	<0.001
Estimated value of selected livestock assets per person	25.4	0.001	90.7	<0.001	47.0	<0.001
Land tenure	6.7	0.035	8.6	0.013	12.4	0.082
Quality of house wall material	43.9	<0.001	163.0	<0.001	158.1	<0.001
Quality of toilet facility	25.2	<0.001	180.6	<0.001	57.8	<0.001
Household Electricity	8.1	0.017	1.8	0.398	0.2	0.890
Accessibility of water supply	8.8	0.067	19.2	0.01	16.9	0.02
Density of household living conditions	13.2	0.001	158.7	<0.001	13.7	0.001
Proportion of household members receiving a regular income	68.1	<0.001	50.2	<0.001	117.1	<0.001
Educational level of household head	20.2	<0.001	127.4	<0.001	75.3	<0.001
Percentage of household members working age adults	3.5	0.177	3.2	0.203	7.6	0.022
Gender of household head	65.4	<0.001	176.6	<0.001	325.7	<0.001
Second most important household income	9.4	0.009	9.7	0.008	30.8	<0.001
Regularity of household having a meal consisting of mielie meal alone, bread alone or worse	32.8	0.007	379.1	<0.001	320.7	<0.001
Car ownership	69.5	<0.001	207.8	<0.001	244.6	<0.001
School attendance score	33.2	0.007	44.5	<0.001	62.6	<0.001

italics denote variables that were considered for but not included in the final Principal Components Analysis. Underlined variables are those that were not considered for inclusion in the PCA analysis. N's for each association vary from 978-1003 dependent on missing data.

The value of non-livestock assets held by the household was strongly associated with the index of household wealth generated by all three methods at a significance level of $p<0.001$. The livestock asset value was also significantly associated with all three indices of household wealth, although each of these associations was less strong than with the previous marker. Land tenure was widely distributed, and was weakly associated with each measure of household wealth, at a significance level of $0.01<p<0.05$. Having a dwelling constructed of better materials was more common among households ranked of higher wealth by all three techniques and the toilet facility indicator was also significantly associated with all three indices. Household electrification was not significantly associated with the measure of household wealth generated by either of the methods that used the survey data, although it was significantly associated with the PWR only index ($p=0.017$). Water accessibility was significantly but weakly associated with both survey-data based measures of household wealth at a significance level of $0.001<p<0.01$, but not with the PWR measure. Density of household living conditions, the proportion of household members receiving an income and the education level and gender of the household head were all significantly associated with all three measures of household wealth ($p<=0.001$). However, the proportion of adults who were of productive age (14-60 years) was not associated with household wealth as estimated by any of the techniques. Households in which the second most important income source was listed was either non-financial or none was listed were more likely to be classified as of low wealth by all three methods. Food security was very strongly correlated with household wealth as measured by both of the survey techniques, and was significantly but less strongly associated with the measure of household wealth generated from the PWR data alone. Similarly, the score denoting attendance and achievement at school of the children living within each household, weighted as described in Chapter 4, was very strongly associated with both of the survey based indices of household wealth, and significantly but less strongly associated with the PWR measure. Finally, car ownership was strongly associated with all three measure of household wealth at a significance level of $p<0.001$.

It was expected that the associations between individual indicators and the survey wealth indices would be generally stronger than those with the PWR index, since the survey wealth indices were calculated from these data. There were, however, striking similarities in the patterns of association seen. The strongest associations between individual indicators and the PWR wealth index were almost all seen for indicators associated at a significance level of $p<0.001$ with both survey indices, while the weakest associations also generally mapped across all three indices. The only exceptions to this were with the indicators of household electrification and water supply.

The final stage of the analysis was to compare the three techniques in their ranking of household wealth. The two survey data methods were quite strongly correlated (Spearman rho = 0.76, $p<0.001$, n=978), and there was a reasonable degree of agreement in their placing of households into wealth rank terciles (Kappa=0.46). Since these indices were partially developed from the same data it was not surprising to see that these data were well correlated. However, it should be noted that the kappa statistic represented an important level of disagreement in the two techniques ranking of household wealth. The PWR wealth index was significantly, but weakly, correlated with both method 2 (Spearman rho = 0.38, $p<0.001$, $n=989$) and method 3 (Spearman rho = 0.35, $p<0.001$, $n=991$). The levels of agreement in placing households into wealth terciles were relatively low (kappa statistics of 0.19 and 0.16 respectively).

A single index of wealth was required for the remainder of the study. Since the three methodologies showed limited agreement in their ranking of household wealth it was decided to choose one rather than combine them. The decision of which method to use could not be made on the basis of comparison with a gold standard measure of economic status (such as income) since such data were not available. None of the three methods described here could be considered *a priori* better than the others. This decision was therefore to some degree subjective. The PWR wealth index was chosen to be used in further analyses. This index had a number of advantages. Firstly, the technique directly allowed households to be grouped into meaningful welfare rankings on the basis of local perceptions of poverty. This was not possible with either of the survey data methods, which might have been divided into terciles or quintiles for the purposes of analysis. Secondly, the PWR data was collected independently from data on socioeconomic roles and so could be used to measure the strength of association between household wealth and the propensity of young people to be in these roles without introducing bias. Method 2 directly used attendance at school of young people in its estimation of household wealth, and so would not be appropriate to use to examine the independent association between these factors. Thirdly, data collected in this study suggested a high degree of internal validity in the ranking of household wealth by independent groups of PWR participants, and it was not possible to generate evidence of this type for either of the other two methods. Fourthly, it was possible to calculate an index of household wealth for a greater number of households using the PWR data. Finally, since differences in welfare levels between households within the study site were likely to be relatively small compared to differences in welfare levels between the site and other areas in the country, it was felt that local perceptions of wealth might be particularly useful for accurately delineating households into wealth groups. Thus, the decision was made to use the PWR household wealth index for the main study analyses. However, in Chapter 9, the sensitivity of the main study findings to this choice of indicator are investigated. Reasons for the differential ranking of household wealth by the three methods are discussed in more detail in Chapter 10.

When the PWR measure of household wealth was applied to the individual level data, there were 272 males classified as from "very poor" households, 535 from households classified as "poor, but a bit better off" and 109 males from households that were "doing OK". Among females the respective figures were 296, 556 and 151 individuals. Before proceeding to the next stage of the analysis the association between household wealth and each of the potential confounders was examined. Among males there was little evidence of an association between household wealth and language group or the number of other 14-25 year olds in the household. However, there was some evidence of an association such that those from wealthier households were least likely to have been born outside the current village of residence (Chi2 = 5.23, p=0.073), and strong evidence of an association between household wealth and orphanhood, with a trend towards those in lower wealth groups being least likely to have both parents still alive (Chi2 = 30.9, p<0.001). Among females, the pattern was similar except that additionally there was some evidence that females from poorer households were less likely to live in households where there were a higher number of other 14-25 year olds.

6.03.3 Individual socioeconomic roles

Table 6.6 shows the distribution of the three major socioeconomic role variables across the age range.

(a) Males

Among the males, 629/917 (68.6%) of those interviewed were attending school or college. There was a steadily linear increase across the full age range in the proportion of individuals who were not attending school, varying from 2.6% among 14-16 year olds to 85.5% among 23-25 year olds. Male school attendance varied from 62.7% to 75.0% by village and there was no significant association with this.

Table 6.6: Major socioeconomic roles of 14-25 year old young people in rural South Africa

	Age (years)								P (between age groups)
	14-16		17-19		20-22		23-25		
MALES (N)	303		274		201		138		
Attending School	**295**	**97.4**	**221**	**80.7**	**92**	**45.8**	**20**	**14.5**	
Not attending school	**8**	**2.6**	**53**	**19.3**	**109**	**54.2**	**118**	**85.5**	
Self employed or domestic worker	0	0.0	2	3.8	5	4.6	6	5.1	
Salary or regular contract work	0	0.0	7	13.2	13	11.9	29	24.6	
Irregular work	0	0.0	9	17.0	12	11.0	14	11.9	
Unemployed	5	62.5	31	58.5	76	69.7	66	55.9	
Unwilling or unable to work	3	37.5	4	7.6	3	2.8	3	2.5	0.001
Non-migrant	**292**	**96.4**	**243**	**88.7**	**159**	**79.1**	**94**	**68.1**	
Temporary migrant	**11**	**3.6**	**31**	**11.3**	**42**	**20.9**	**44**	**31.9**	
Not currently sleeping at home*	8	72.7	25	80.6	29	69.0	30	68.2	0.649
Not sleeping at home for more than 6months of previous year*	6	54.6	15	50.0	30	71.4	27	61.4	0.302
Current highest level of educational attainment									
None	0	0.0	2	0.7	0.0	0.0	1	0.7	
Attended primary only	140	46.2	55	20.2	19	69.5	14	10.1	
Attended secondary school	163	53.8	209	76.6	140	69.7	82	59.4	
Completed secondary school	0	0.0	4	1.5	28	13.9	32	23.2	
Attended university / technikon	0	0.0	3	1.1	14	7.0	9	6.5	<0.001
FEMALES (N)									
In School	**303**	**93.8**	**228**	**76.5**	**72**	**30.4**	**20**	**13.8**	
Not attending school	**20**	**6.2**	**70**	**23.9**	**165**	**69.6**	**125**	**86.2**	**<0.001**
Self employed or domestic worker	0	0.0	3	4.3	4	2.4	8	6.4	
Salary or regular contract work	0	0.0	1	1.4	10	6.1	14	11.2	
Irregular work	0	0.0	3	4.3	6	3.6	14	11.2	
Unemployed	14	70.0	61	87.1	143	86.7	88	70.4	
Unwilling or unable to work	6	30.0	2	2.9	2	1.2	1	0.8	<0.001
Non-migrant	**305**	**94.4**	**274**	**92.0**	**201**	**84.8**	**116**	**80.0**	
Temporary migrant	**18**	**5.6**	**24**	**8.0**	**36**	**15.2**	**29**	**20.0**	**<0.001**
Not currently sleeping at home*	16	88.9	16	66.7	27	75.0	21	72.4	0.420
Not sleeping at home for more than 6months of previous year*	10	55.6	13	54.2	20	55.6	17	58.6	0.990
Current highest level of educational attainment									
None	3	6.9	0	0.0	2	0.8	0	0.0	
Attended primary only	109	33.4	44	14.8	18	7.6	16	11.0	
Attended secondary school	209	64.7	243	81.5	172	72.6	93	64.1	
Completed secondary school	1	0.3	10	3.4	35	14.8	22	15.2	
Attended university / technikon	1	0.3	1	0.3	10	4.2	14	9.7	<0.001

* These categories are not mutually exclusive

The remaining 288 individuals were not attending school or college, and the proportion in this group increased with age. Of these individuals the majority were unemployed and looking for a job for the last year (n=178, 61.8%), while a further group had accessed irregular work over the past year (n=35, 12.2%). A small number of individuals were self employed (n=13, 4.5%) or described themselves as unwilling or unable to work (n=13, 4.5%). Some 49 individuals had regular contract or salaried employment during the past year (17.0%). The proportion of those not attending school who were in this category increased from 0 (0.0%) among the 14-16 year age group to 29/118 (24.6%) among 23-25 year olds.

Level of education attained also varied strongly with age. Only three males in the whole sample reported that they had not attended school at all. Among the youngest age group some 141 individuals (46.4%) reported that they had only attended primary school, while this figure was lower among the older age groups. Most individuals had attended secondary school. However, the proportion completing secondary school, including those who had also attended university or technikon was low, increasing from 0 (0.0%) of those in the youngest age group to just 41/138 (29.7%) of those in the oldest age group. After adjustment for age using Mantel Haenszel techniques there was no significant association between current attendance at school or college and the likelihood of having completed secondary school education.

Temporary migrancy was defined as a composite of the variables relating to currently sleeping in the home and number of months sleeping away from the home during the past year. Among the males, 128 individuals (14.0%) were classified as temporary migrants. Similar numbers of individuals were classified as migrants by both of the definitions used. However, the two definitions did not agree in all cases. For example, of males classified as migrant by the "current" definition, only 40/88 (45.5%) were also classified as migrant by the "whole year" definition. Temporary migrancy was relatively uncommon among those in the youngest age groups (Table 6.6). Among males, the figure rose steadily over the age range to 31.9% among 23-25 year olds.

Finally, the association between the two socioeconomic roles was investigated among the males. Since both roles were strongly associated with age, Mantel Haenszel techniques were used to adjust the analysis for this factor. Mantel Haenszel techniques were also used to examine whether the association between school attendance and temporary migrancy was the same among those aged 14-19 years and 20-25 years, thus checking for effect modification. The two socioeconomic role variables were statistically significantly associated among the younger but not the older group (Mantel-Haenszel test for homogeneity, p=0.0004). Among the younger age group, those attending school were much less likely to be migrants (OR 0.13, 95% CI 0.06-0.27, Chi2 42.9, p<0.0001), while among the older group the association was only of borderline statistical significance (OR 0.62, 95% CI 0.30-1.08, Chi2 2.89, p=0.089).

(b) Females

Slightly less females than males were attending school (625/1005, 62.2%), although this group were slightly older than the males. The pattern of school attendance by age among the females was similar to that of the males at the beginning and end of the age period. However, among females the age trend was less linear, with a suggestion of decreased school enrolment relative to males among those aged 17-23. Female school enrolment varied significantly by village from 46.5% to 74.2% (p=0.001).

Some 380 females were not attending schools, and a higher proportion of those were unemployed than for the males (n=306, 80.5%). Relatively small numbers of individuals were self employed (n=15), had irregular work (n=23) or were unwilling or unable to work (n=11) during the past year. Only 25 females (6.6%) were in regular contract or salaried employment, and this increased over the age range from 0 (0.0%) to 14 (11.2%).

Slightly less females than males reported only attending primary school (n=187, 18.6%) and this was again concentrated among the youngest group. A similar proportion had completed secondary school (n=68, 6.8%) or attended technikon / university (n=26, 2.6%).

Among females, the association between schooling and migration also varied with age (Mantel Haenszel test for homogeneity, p=0.003). Among the youngest age group (14-16 years) there was only one individual who was a temporary migrant and out of school so statistical techniques to estimate the association between the two were low in this strata. However, among those aged 17-19 years, temporary migrancy was much more common among those out of school, as for the males (13/70, 18.6% of those not attending school were migrants, compared to 11/228, 4.8% of those attending school, p<0.001). There was no association between temporary migrancy and attendance at school among either of the older age groups.

Finally, before going on to examine the association between household wealth and individual socioeconomic roles, the crude association between these roles and potential confounders was also examined since this information would be used to inform the control of confounding for analyses presented later in the thesis. The majority of the population of both sexes had Sepedi as first language, and this factor was not associated with either socioeconomic role among either sex (Table 6.7). Coming from a household with more than two 14-25 year olds was positively associated with being out of school among both sexes, and additionally among males was significantly associated with temporary migration. Finally, having lost either parent was significantly associated with non-attendance at school among both sexes, and was also positively associated with temporary migrancy among males.

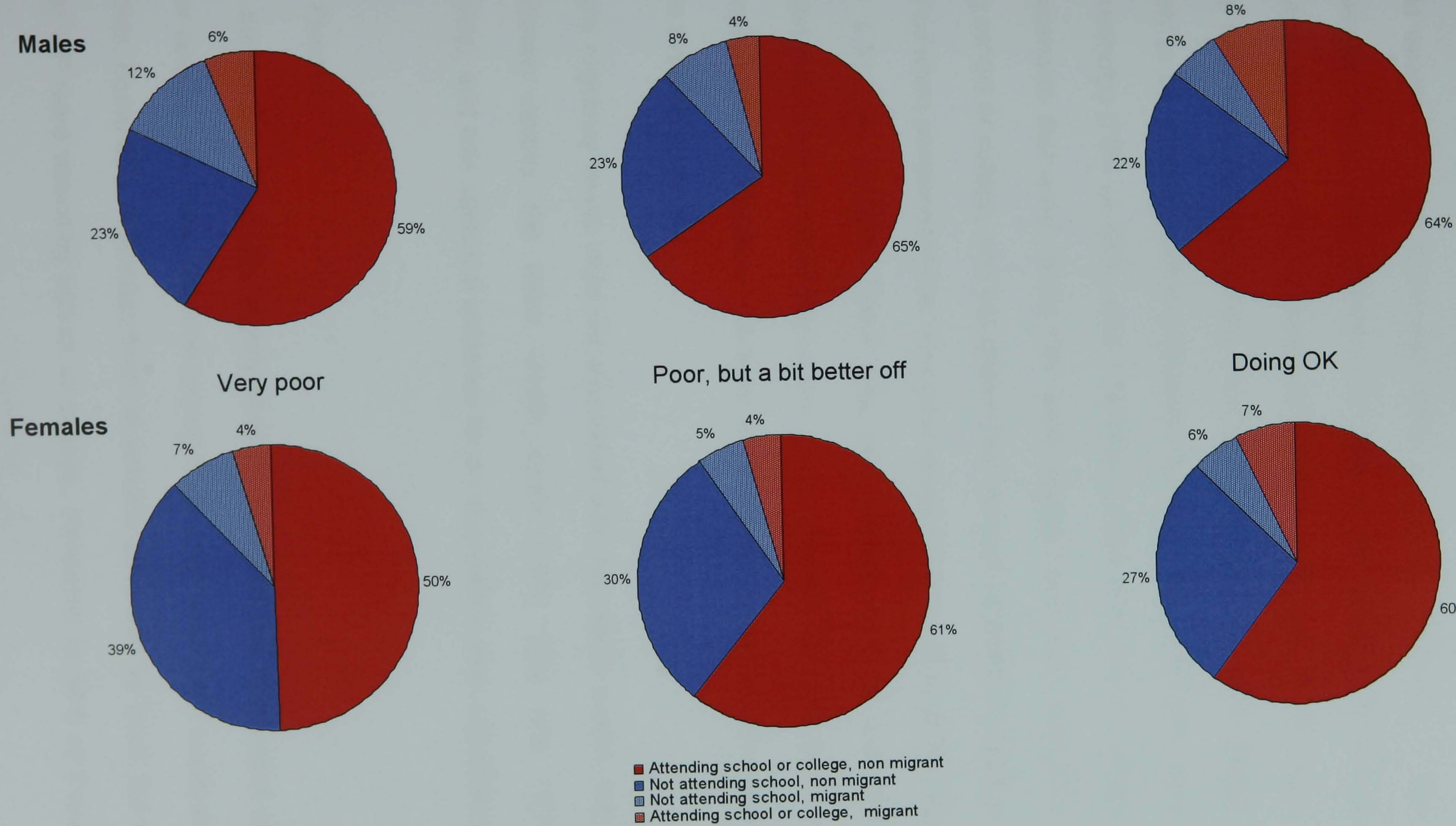
6.03.4 The association between household wealth and individual socioeconomic roles

Figure 6.8 shows pie-charts depicting the association between household wealth and the socioeconomic roles under examination in this thesis for males. In the figures, the red sections of the pie charts represent those attending school or college, while the blue sections represent those not doing so. Overlaid on this are patterns of temporary migrancy, with numbers of temporary migrants identified by the dotted sections of the pie chart, and non-migrants identified by the solid sections.

Table 6.7: The relationship between potential confounders and livelihood roles among males and females aged 14-25

Potential Confounder	Out of school		In School		P	Not temporary Migrant		Temporary migrant		P
Males										
N	288		628			788		128		
	n	%	n	%		n	%	n	%	
Sepedi speaking	258	89.6	574	91.4	0.376	720	91.4	112	87.5	0.159
Born away from village of current residence	77	26.7	138	22.1	0.114	186	23.6	29	22.7	0.814
More than 2 young people (14-25 in the home)	184	63.9	351	55.9	0.023	446	56.6	89	69.5	0.006
Orphan	123	42.7	195	31.1	0.001	262	33.3	56	43.8	0.021
Females										
N	380		623			896		107		
Sepedi speaking	340	89.5	573	92.0	0.179	814	90.9	99	92.5	0.567
Born away from village of current residence	107	28.2	156	25.0	0.276	229	25.6	34	31.8	0.167
More than 2 young people (14-25 in the home)	260	68.4	345	55.4	<0.001	541	60.4	64	59.8	0.910
Orphan	157	41.3	165	26.5	<0.001	284	31.7	38	35.5	0.424

Figure 6.8: Pie charts showing the association between household wealth and socioeconomic role variables among males and females in Sekhukhuneland villages in 2001



(a) Males

Among the males, attendance at school was not significantly associated with household wealth, although there was some suggestion that those in the highest wealth bracket were slightly more likely to be attending school or college. After Mantel-Haenszel adjustment for age, in two age bands, those who were from households that were "poor, but a bit better off" were slightly more likely to be attending school compared to those who were from "very poor" households (70% vs. 65%, aOR 1.32 95% CI 0.90-1.94, p=0.16). Those from households that were "doing OK" were slightly more likely still to be attending school or college, and this relationship reached borderline statistical significance when compared to the "very poor" group (72% vs. 65%, aOR 1.80 95% CI 0.91-3.55, p=0.07). There was some weak evidence that the association was to some extent dependent on age (Mantel Haenszel test for homogeneity p=0.10). However, numbers of individuals in some cells were low in the stratified analysis.

Temporary migrancy levels were not associated with household wealth, with little variation across the three wealth strata (18%, 12% and 15% respectively), and little statistical evidence for an association after adjustment for age.

(b) Females

Among the females there was slightly stronger evidence that attendance at school or college was more common among those who were from wealthier households. Some 65% of females from households described as "poor, but a bit better off" were attending school or college, compared to 54% of those from "very poor" households (Mantel Haenszel age-adjusted OR 1.40 95% CI 0.96-2.03, p=0.08), while this effect was stronger when considering those from households that were "doing OK" (67% vs. 54%, aOR 1.72 95% CI 1.03-2.87, p=0.04). There was no evidence to suggest that the association between household wealth and school attendance was dependent on age.

Similarly to the males, there was no association between temporary migrancy and household wealth, with little variation across the three wealth strata (11%, 9% and 13% of females were temporary migrants respectively, see Figure 6.8).

6.04 Summary of key findings

Households in the study sample experienced a range of wealth from fragile access to basic needs among the poorest group to secure access to jobs, housing and schooling among those at the top end of the scale. Household income was generated through local and migrant labour, although self employment and access to pensions were also important. Many households additionally relied on support from their neighbours to some degree.

Three measures of household wealth were generated from different techniques. Though the techniques appeared generally to agree in which dimensions of welfare were associated with relative household wealth, there was not substantial agreement in the ranking of household wealth between the survey based methods and PWR based method. The indicator of household wealth developed from PWR data was selected for use in further analyses. This indicator showed a high level of internal consistency, was effective in ranking households on the basis of locally important markers of poverty, and could be used to divide households into those that were very poor (36.8%), poor, but a bit better off (56.2%) and doing OK (13.0%).

Individuals aged 14-25 years were increasingly likely not to be attending school, to have completed secondary education and to be temporary migrants over the age range. Most of those who left school remained unemployed, although levels of secure employment increased with age. Among both sexes at younger ages, migrants were more likely to be out of school, but at older ages temporary migrants were just as likely to be in school or college as to not be doing so. Attendance at school was high at the low end of the age range in all groups, but was statistically significantly associated with household wealth among females, while a borderline association was noted among males. Temporary migrancy was not associated with household wealth among either sex.

Chapter 7. HIV awareness, sexual behaviour and HIV infection

7.01 Introduction

The outcome measures under consideration in this thesis were indicators of the risk of HIV infection. These measures included dimensions of HIV awareness reported to be associated with risk behaviours in previous research, characteristics of sexual behaviour reported in epidemiological and other studies to affect risk of HIV acquisition, and the presence of antibodies to HIV indicative of prior acquisition of the virus. This chapter describes how these indicators were distributed within the study sample and explores how they were associated with each other.

7.02 Objectives of chapter

1) Describe the distribution of a variety of HIV awareness variables among males and females
2) Describe the distribution of individual-level sexual behaviour characteristics among males and females
3) Describe the distribution of partnership-level sexual behaviour characteristics reported by males and females who had been sexually active during the previous year
4) Describe the distribution of prevalent HIV infection in the population and assess the association between socio-demographic, HIV awareness and sexual behaviour factors and HIV infection among males and females in the study sample.

7.03 Results

7.03.1 HIV awareness variables

Among both males and females the proportion of individuals not able to correctly confirm two true statements about HIV decreased with age (see Table 7.1). Lack of knowledge that a healthy looking person can be infected with HIV was similar among males and females in both age groups (57.8% vs. 43.0% at 14-19 years, and 56.3% vs. 55.4% for 20-25 years). Knowledge that a pregnant woman can transmit HIV to her unborn baby was more common among both sexes. There was no association between sex and the variable "lack of HIV knowledge" after adjustment for age (Mantel Haenszel p=0.452).

The proportion of young people who did not communicate with parents / guardians about sex / sexuality during the previous year did not vary to a large degree with age among females, although among males this was slightly more common among the younger group. Conversely, communication with others in the household was relatively rare at the younger age group, but more common among the older age group among both sexes. Overall, females were more likely than males to communicate with someone in their household, after adjustment for age (p=<0.001). Some additional information was also available on other patterns of communication. Both males and females more often reported communication with their peers or with teachers than with parents or household members (data not shown). The nature of such communication was not known.

Males and females were both more likely to have had an HIV test at older ages, and this was much more common among females than males. Some additional information was also available on this factor. Among the 48 males who had had a test and more information was available, 39 (81.3%) reported that it had been voluntary, whereas only 52/94 (55.3%) of females reported this. Some 10/48 (20.8%) of males and 13/94 (13.8%) of females reported that they did not receive the result of the test.

Stigmatising attitudes were widely held throughout the study population of both males and females. More younger individuals than older individuals held each of the stigmatising attitudes enquired about among either sex. However, there was little association between sex and the prevalence of stigmatising attitudes after adjustment for age.

Table 7.1: The age distribution of indicators of HIV awareness among males and females aged 14-25 in rural South Africa

	Males					Females				
	Age					Age				
	14-19		20-25		P	14-19		20-25		P
	n	(%)	n	(%)		n	(%)	n	(%)	
N	577		339			621		382		
Cannot correctly identify that ..										
A healthy looking person can be infected with HIV	335	58.1	149	44.0	<0.001	353	56.8	171	44.8	<0.001
A pregnant woman infected with HIV can transmit the virus to her unborn child	124	21.5	37	11.0	<0.001	101	16.3	44	11.6	0.039
Lack of HIV Knowledge	**364**	**63.1**	**165**	**48.7**	**<0.001**	**374**	**60.2**	**187**	**49.0**	**<0.001**
During the past 12 months has not spoken about sex, sexuality or STIs (including HIV) to;										
Parents / guardians	459	79.6	250	73.8	0.043	444	71.5	266	69.6	0.528
Other household members	451	78.2	195	57.5	<0.001	438	70.5	235	61.7	0.004
Lack of household communication	**385**	**66.7**	**163**	**48.1**	**<0.001**	**621**	**55.2**	**178**	**46.6**	**0.008**
Has not had an HIV test	**562**	**97.4**	**306**	**90.3**	**<0.001**	**588**	**94.7**	**321**	**84.0**	**<0.001**
Not or don't know whether willing to ...										
Eat with someone known to be HIV positive	326	56.7	120	35.4	<0.001	340	54.8	175	45.8	0.006
Look after a relative known to be HIV positive	85	14.7	24	7.1	0.001	100	16.1	42	11.0	0.024
Buy food from someone known to be HIV positive	304	52.7	120	35.5	<0.001	313	50.4	161	42.2	0.011
Holds Stigmatising attitudes	**386**	**66.9**	**155**	**45.7**	**<0.001**	**405**	**65.2**	**215**	**56.3**	**0.005**

There were missing values for some variables; Communication with sexual partner (3), with teacher (2), perception of risk (18).

7.03.2 Individual-level sexual behaviour characteristics

Tables 7.2 and 7.3 provide summaries of sexual behaviour characteristics of males and females respectively throughout the age distribution under study.

(a) Males

Among males, as age increased the proportion of individuals who reported they had never had sex decreased from 71.6% at 14-16 years to 5.8% by 23-25 years. Among those who had ever been sexually active, a similar proportion at each age group reported that they had not had any sexual partners during the previous year (between 4.6% to 7.5% of all individuals). Of those currently sexually active, the proportion of individuals having a single non-spousal partner was approximately double that reporting multiple partners at each age group, with both of these indicators rising in prevalence with age.

As would be expected the number of sexual partners during the lifetime rose with age, rising to its highest level among those aged 23-25 years (median 5, interquartile range 3-10). The number of sexual partners reported in the previous year was most often a single partner at all ages except in the youngest age group. There was a strong association between the total number of lifetime partners and the number of partners reported for the last year among males.

Median age at sexual intercourse from a survival analysis was 17 years among the two older age groups, among who most individuals were sexually active. A log rank test did not suggest any difference in age of sexual debut by age group (see Table 7.2). Most males reported that they wanted to have sex the first time. A decreasing proportion with age reported that sex had "just happened" the first time. Only very few males reported being forced into sexual debut by a partner, and most of these were among the youngest group.

At the youngest age group, all males reported that all of their sexual partners during the past year were of a similar age to themselves (within 3 years). However, as age increased, a larger proportion of males reported that all of their partners were at least 3 years younger than themselves so that this represented 59% of all males aged 23-25 years. Only a small proportion of males reported any partnerships with females who were older than themselves by more than 3 years. Overall 180/916 (19.7%) of males reported any sexual partnerships where only they provided resources. Younger males were least likely to report this although this was not statistically significant.

During the previous year, a majority of males reported that they had never used a condom with any sexual partner at all ages, with the highest figure reported among the youngest males (68.1%), although this was also not associated with age.

Table 7.2: Selected individual sexual history characteristics among males of different ages in rural South Africa

		Age (years)								P (between age groups)
		14-16		17-19		20-22		23-25		
		n	%	n	%	n	%	n	%	
N		***303***		***274***		***201***		***138***		
Sexual activity summary	Never had sex	217	71.6	88	32.1	25	12.4	8	5.8	<0.001
	Secondary abstinence during last year	14	4.6	17	6.2	15	7.5	9	6.5	
	Sex with a single non-spousal partner	52	17.2	115	42.0	109	54.2	83	60.1	
	Sex with multiple partners	20	6.6	54	19.7	52	25.9	38	27.5	
Age of sexual debut	Median (IQ range)	- (15,-)		16 (15,-)		17 (16,18)		17(15,18)		0.667
Experience at sexual debut[a] (%)	Wanted	62	72.1	141	75.8	144	81.8	114	87.7	0.005
	Just happened	20	23.3	43	23.1	31	17.6	16	12.3	
	Forced	4	4.7	2	1.1	1	0.6	0	0.0	
No. of sexual partners in last year	Median (IQ Range)	0 (0,0)		1 (0,1)		1 (1,2)		1 (1,2)		0.001
No. of lifetime sexual partners	Median (IQ Range)	0 (0,1)		2 (0,4)		3 (2,6)		5 (3,10)		0.001
N (Sexually active in last year)		***72***		***169***		***161***		***121***		
Frequency of intercourse[b]	All partners sex more than 5 times in past year	25	34.7	80	47.3	81	50.3	81	66.9	<0.001
	Any partners <5times sex in past year	47	65.3	89	52.7	80	49.7	40	33.1	
Age difference of partners during past year[b]	All partners at least 3 years younger	0	0.0	11	6.8	51	33.8	69	59.0	<0.001
	At least one partner of similar age	70	100.0	150	93.2	100	66.2	44	37.6	
	At least one partner >3yrs older	0	0.0	0	0.0	0	0.0	4	3.4	
Resource exchange between partners in last year[b]	All partnerships no, reciprocal or female to male resource exchange	53	73.6	117	69.2	95	59.0	78	64.5	0.102
	At least one partnership male provides resources but female does not	19	26.4	52	30.8	66	41.0	43	35.5	
Condom use during last year[b]	Never used a condom with any partner	49	68.1	86	50.9	83	51.6	63	52.0	0.285
	Use a condom sometimes with all or some partners	16	22.2	56	33.1	51	31.7	41	33.9	
	Always / nearly always use a condom with all partners	7	9.7	27	16.0	27	16.8	17	14.1	

[1] Chi-Square, [2] Kruskal-Wallis, [3] Log Rank Test. [a] Among those ever sexually active only [b] Among those who reported a non-spousal partnership during the previous year.
There were missing values for some variables; age at sexual debut (15) used in log-rank test only, age difference with partners (24).

(b) Females

Among females, the proportion of those who had never had sex fell more steeply between the ages of 14-16 and 17-19 years than for males. Similarly to males, most females at the older end of the age spectrum had been sexually active. As with males, secondary abstinence during the previous year was reported by a minority (2.8% to 10.6% among the age groups). The proportion of females who reported a single sexual partner during the previous year rose rapidly between the ages of 14-16 and 17-19, and then rose less quickly with age. A much smaller proportion of females than males reported multiple sexual partners during the past year.

As for males, and as expected, a higher number of lifetime sexual partners were reported by individuals at older ages. The numbers of reported lifetime sexual partners was lower for females (median 3, IQ range 2-4 at 23-25 years) than it was for males. Similarly to the males, most individuals reported a single sexual partner during the past year in all but the youngest age group. Also similarly to the males, there was a strong association between the total number of lifetime partners and the number of partners reported for the last year.

Examining the Kaplan-Meier curves (not shown) suggested that slightly more males than females tended to report sexual activity at very young ages (under 14), but by the late teens cumulatively more females had begun to be sexually active. Among females there was some evidence to suggest that rates of new sexual debut were lower among younger women.

Table 7.3: Selected individual sexual history characteristics among females of different ages in rural South Africa

		Age (years)								P (between age groups)
		14-16		17-19		20-22		23-25		
		n	%	n	%	n	%	n	%	
N		***323***		***298***		***237***		***145***		
Sexual activity summary	Never had sex	225	69.7	60	20.1	10	4.2	2	1.4	<0.001
	Secondary abstinence during last year	9	2.8	26	8.7	25	10.6	10	6.9	
	Sex with a single non-spousal partner	77	23.8	191	64.1	188	79.3	123	84.8	
	Sex with multiple partners	12	3.7	21	7.1	14	5.9	10	6.9	
Age of sexual debut	Median (IQ range)	- (15,-)		16 (15,18)		17 (16,18)		17 (16,18)		0.049
Experience at sexual debut[a] (%)	Wanted	44	44.9	107	45.0	132	58.2	77	53.9	0.002
	Just happened	42	42.9	107	45.0	90	39.7	58	40.6	
	Forced	12	12.2	24	10.0	5	2.2	8	5.6	
No. of sexual partners in last year[a]	Median (IQ Range)	0 (0,1)		1 (0,1)		1 (1,1)		1 (1,1)		<0.001
No. of lifetime sexual partners	Median (IQ Range)	0 (0,1)		2 (1,3)		2 (1,4)		3 (2,4)		<0.001
N (Sexually active in last year)		***89***		***212***		***202***		***133***		
Frequency of intercourse[b]	All partners sex more than 5 times in past year	35	39.3	128	60.4	150	74.3	100	75.2	<0.001
	Any partners <5times sex in past year	54	60.7	84	39.6	52	25.7	33	24.8	
Age difference of partners during past year[b]	All partners at least 3 years younger	0	0.0	0	0.0	0	0.0	1	0.8	0.019
	At least one partner of similar age	45	56.3	100	48.5	72	38.1	47	37.3	
	At least one partner >3yrs older	35	43.7	106	51.5	117	61.9	78	61.9	
Resource exchange between partners in last year[b]	All partnerships no, reciprocal or female to male resource exchange	21	23.6	50	23.6	56	27.7	45	33.8	0.173
	At least one partnership male provides resources but female does not	68	76.4	162	76.4	146	72.3	88	66.2	
Condom use during last year[b]	Never used a condom with any partner	39	43.8	109	51.4	97	48.0	68	51.1	0.871
	Use a condom sometimes with all or some partners	34	38.2	76	35.9	75	37.1	45	33.8	
	Always / nearly always use a condom with all partners	16	18.0	27	12.7	30	14.9	20	15.0	

[1] Chi-Square, [2] Kruskal-Wallis, [3] Log Rank Test. [a] Among those ever sexually active only [b] Among those who reported a non-spousal partnership during the previous year. There were missing values for some variables
There were missing values for some variables; age at sexual debut (7), age difference with partners (35).

At all age groups, the proportion of females who wanted to have sex at sexual debut was lower than for males, and the proportion reporting that it just happened or that they were forced was higher. Among the females, report of forced sexual debut was commonest in the youngest age group (12.2%) and least common among those aged 20-22 years (2.2%). A lower proportion of younger than older females reported that they wanted to have sexual intercourse at first intercourse.

The youngest females reported partnerships with individuals who were of the same age and with partners who were more than three years their senior. As the age of females increased so did the proportion of individuals reporting partnerships that were characterised by wide age gaps. Very few females reported sexual partnerships with males at least 3 years younger than themselves. Females more often reported some partnerships characterised by male-to-female only resource flow than did males at all ages. This characteristic was not associated with age among females. Use of condoms was not associated with age, with about half of all females reporting that they never used a condom with any partner during the past year. More detail of the inter-relationship of partnership characteristics is held in the next section.

7.03.3 Partnership-level sexual behaviour characteristics

Some 515 males and 627 females reported details of sexual relationships during the previous year. The males reported a total of 774 sexual partnerships while the females reported 702 sexual partnerships. Details of only the first three partnerships for any single individual were recorded. Consequently, no data was available on 24 partnerships that were reported by males. These partnerships were reported by individuals who had had 7 (n=1), 6 (1), 5 (4) and 4 (9) partners during the previous year and a further two were unreported for other reasons. The total number of partnerships reported by males on which additional data was available was therefore 748. For females, details were not available for only two partnerships, comprising one partnership from each of two individuals who reported a total of four partnerships during the previous year, giving a total of 700 partnerships reported on by females.

(a) Males

Among males, 438 (58.6%) partnerships were reported on in which sexual intercourse had occurred more than 5 times during the previous year, while the remaining 310 (41.4%) partnerships were partnerships where sex had occurred less than five times during the previous year. Males aged 20-25 reported a lower proportion of sexual relationships characterised by a lower frequency of sexual intercourse (144/413, 34.9% of partnerships reported by older males compared to 166/335, 49.6% of those reported by younger males, $p<0.001$).

Partnerships reported by younger males were almost all with females who were of a similar age to the interviewee, while among the older males, reported partnerships were more likely to be with females who were more than three years younger. Age differences between partners did not differ significantly with the regularity of sexual intercourse during the previous year among either age group.

Table 7.4: Partnership Characteristics in last sexual partnership reported by males

		Males									
		14-19 years					20-24 years				
		Partners with whom sexual intercourse more than 5 times in past year		Partners with whom sexual intercourse 5 times or less past year		P (Chi-Square)	Partners with whom sexual intercourse more than 5 times in past year		Partners with whom sexual intercourse 5 times or less past year		P (Chi-Square)
Total number of partnerships		169		166			269		144		
		n	%	n	%		n	%	n	%	
Relative age of partner	Female more than 3 years younger than male	9	5.7	9	5.7		138	54.3	65	49.6	
	Within 3 years	149	94.3	150	94.3		115	45.3	63	48.1	
	Male more than 3 years older than female	0	0.0	0	0.0	0.989	1	0.4	3	2.3	0.173
Resource exchange	Reciprocal	62	36.7	26	15.7		95	35.3	38	26.4	
	Male to female	41	24.3	43	25.9		93	34.6	30	20.8	
	Female to male	24	40.2	17	10.2		31	11.5	23	16.0	
	No resource exchange	42	24.9	80	48.2	<0.001	50	18.6	53	36.8	<0.001
Regularity of Condom Use	Never	79	46.7	92	55.4		111	41.3	73	50.7	
	Less than half the time	38	22.5	32	19.3		57	21.2	27	18.8	
	More than half the time	26	15.4	13	7.8		47	17.5	16	11.1	
	Always or nearly always	26	15.4	29	17.5	0.113	54	20.1	28	19.4	0.199
Condom Use at Last Sex	Yes	47	27.8	56	33.7		92	34.2	52	36.1	
	No	122	72.2	110	66.3	0.240	177	65.8	92	63.9	0.698

There were missing values for some variables; relative age of partner (45).

Among partnerships reported by younger males, those partnerships where sex had occurred relatively few times (<5) during the previous year were least likely to have involved any resource exchange between partners (48.2% reporting no resource exchange). A quarter of these partnerships were reported by younger males as involving only the male providing some resources to the female partner (25.9%). In non-spousal partnerships where sex during the past year was more frequent, reciprocal resource exchange was the most common type of partnership noted (36.7%). Among partnerships reported by the older group of males, a lower proportion than those reported by younger involved no resource exchange, regardless of frequency of sexual intercourse (18.6% vs. 24.9% for partnerships where sex occurred more than 5 times, 36.8% vs. 48.2% for partnerships with less frequent intercourse). Resource exchange patterns were significantly associated with the frequency of sexual intercourse during the past year at both age groups.

Never having used a condom during the past year was the most frequently reported characteristic of partnerships of either type reported by males. Among the younger males, this was more often reported for partnerships involving less frequent sexual contact (55.4% vs. 46.7%), though this did not reach statistical significance. Among partnerships reported by the older men, never having used a condom was again more frequently reported for partnerships where sex had occurred less than 5 times (50.7% vs. 41.3%), although the Chi-square test comparing the distribution across all partnership types was not significant (p=0.199). Condom use at last sex was reported for 247/748 (33.0%) of all partnerships and was not significantly different across partnership types at either age group.

Finally for males, it was investigated whether sexual frequency, age differences and resource exchange patterns in relationships were related to condom use at last sex using regression models. In a model comprising all non-spousal partnerships with condom use at last sex as the outcome variable, and including independent terms for age, age difference, resource exchange and frequency of sex in the past year, none of these factors was found to be significantly associated.

(b) Females

As for the males, older females less often reported partnerships with lower sexual frequency (87/361, 24.1%) than did younger women (149/339, 44.0%, $p<0.001$). Among partnerships reported by younger females, relationships where sex had occurred more frequently were more likely to be characterised by a wider age gap than those where sex had been less frequent (Table 7.5). Among partnerships reported by the older group this association was not seen, although the proportion of relationships with a wider age gap was larger than among the younger group for both partnership types.

Table 7.5: Partnership Characteristics in last sexual partnership reported by females

		Females									
		14-19 years					20-24 years				
		Partners with whom sexual intercourse more than 5 times in past year		Partners with whom sexual intercourse 5 times or less past year		P (Chi-Square)	Partners with whom sexual intercourse more than 5 times in past year		Partners with whom sexual intercourse 5 times or less past year		P (Chi-Square)
Total number of partnerships		190		149			274		87		
		n	%	n	%		n	%	N	%	
Relative age of partner	Female more than 3 years younger than male	0	0.0	0	0.0		0	0.0	1	1.3	
	Within 3 years	80	44.2	90	64.3		99	38.2	32	41.6	
	Male more than 3 years older than female	101	55.8	50	35.7	<0.001	160	61.8	44	57.1	0.154
Resource exchange	Reciprocal	20	10.5	13	8.7		24	8.8	8	9.2	
	Male to female	146	76.8	108	72.5		194	70.8	52	59.8	
	Female to male	9	4.7	7	4.7		22	8.0	8	9.2	
	No resource exchange	15	7.9	21	14.1	0.319	34	12.4	19	21.8	0.151
Regularity of Condom Use	Never	94	49.5	70	47.0		129	47.1	46	52.9	
	Less than half the time	39	20.5	42	28.2		53	19.3	22	25.3	
	More than half the time	28	14.7	14	9.4		49	17.9	6	6.9	
	Always or nearly always	29	15.3	23	15.4	0.253	43	15.7	13	14.9	0.077
Condom Use at Last Sex	Yes	53	27.9	65	43.6		67	24.5	21	24.1	
	No	137	72.1	84	56.4	0.003	207	75.6	66	75.9	0.953

There were missing values for some variables; relative age of partner (42).

Younger females most often reported that sexual partnerships were characterised by resources being provided by the male, but not the female partner. Some 36/339 (10.6%) of partnerships were described as characterised by no resource flow. This pattern was very similar to that seen among partnerships reported by the older age group. There was apparently no association between resource exchange patterns and frequency of sexual intercourse in the past year at either age group for sexual partnerships reported by females.

Regularity of condom use was not significantly associated with the frequency of sexual intercourse among either group, although in the older age group the relationship noted was of borderline significance in that condom use was reported never or less than half the time more often for relationships in which sex occurred less than five times during the past year. However, condom use at last sex was significantly less likely to be reported for partnerships where sex had occurred more frequently in the previous year among partnerships reported by 14-19 year olds. There was no significant relationship in this association among the older group.

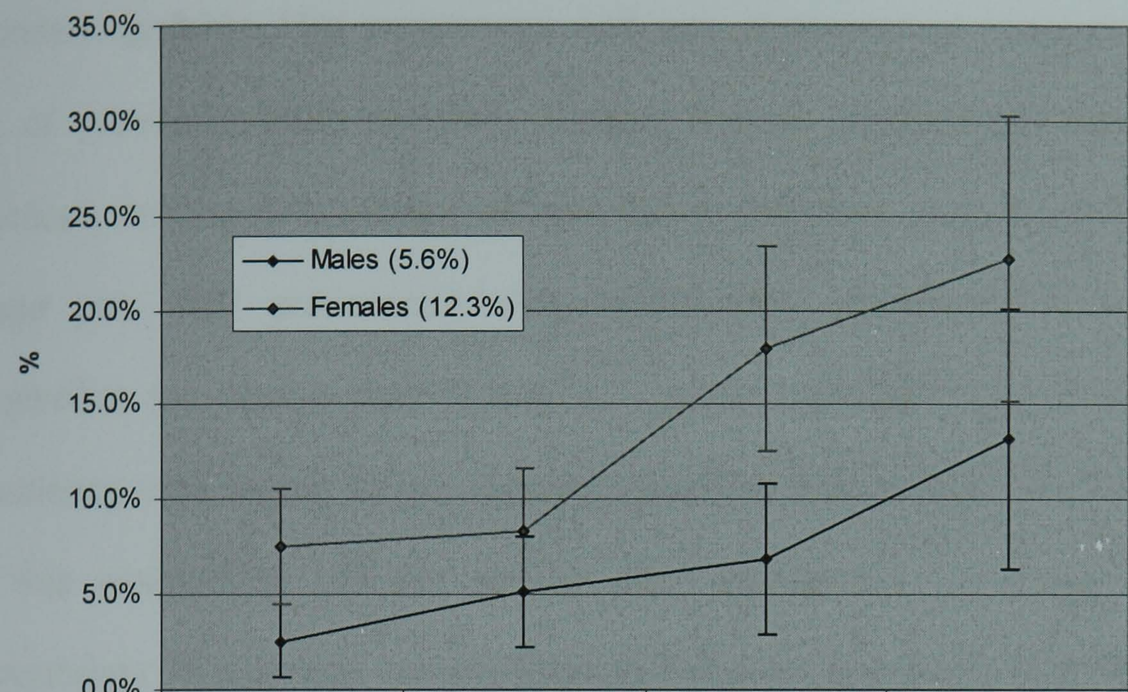
Among females there was some suggestion from bivariate analyses that among the whole age range, partnerships where sex had occurred less frequently were less likely to involve male to female only resource exchange than other resource exchange types. These partnerships were also less likely to be characterised by a wide age gap and more likely to have involved condom use at last sex. These associations persisted after adjustment for age. Partnerships characterised by a wider age gap were less likely to be associated with condom use at last sex and more likely to be associated with a male to female only resource exchange pattern. Male to female resource exchange was significantly more likely to be associated with condom use at last sex than all other resource exchange partnership types. In a model with condom use at last sex as the dependent variable and including terms for all of the other partnership level characteristics as well as age, the association between age difference and condom use suggested that partnerships where the age difference was greatest were more likely not to involve condom use although this was no longer significant. Some 66.7% of partnerships where partners were within 3 years of age reported no condom use at last sex, while 74.1% of partnerships where the male was more than 3 years older reported this (aOR adjusting for all other partnership characteristics 1.34 95% CI 0.94-1.90). However, it was still noted that not using a condom at last sex was more frequently reported for partnerships characterised by more frequent sexual intercourse (aOR 1.46 95% CI 1.01-2.11) and less common among those involving male-to-female only resource exchange (aOR 0.63 95% CI 0.42-0.94).

7.03.4 Distribution of prevalent HIV infection

Data on HIV infection were available for 763 males and 880 females. Overall 5.6% (43/763) of males and 12.3% (108/880) of females were HIV infected. Among males there was one village in which no HIV infections were detected, while the highest prevalence in a single village was 18/96 (18.8%). Among females the range was from 8.0% (9/112) to 20.8% (25/120). The highest and lowest HIV prevalence occurred in the same villages for both males and females. HIV prevalence varied significantly by village for males ($p<0.001$)

even after exclusion of the village in which no HIV infections were detected and with borderline significance for females (p=0.08). Among both sexes, the lowest infection rates were seen among those of the youngest age group. Among both sexes, prevalence of infection rose approximately linearly with age (Figure 7.6). Females were significantly more likely to be HIV infected than males after adjustment for age (Mantel Haenszel age-adjusted OR 2.31, 95% CI 1.59-3.36).

Figure 7.6: Estimated prevalence of HIV infection and 95% confidence intervals by age and sex among the study sample



	14-16yrs	17-19yrs	20-22yrs	23-25yrs
Males (5.6%)	2.6%	5.2%	6.8%	13.3%
Females (12.3%)	7.5%	8.3%	18.0%	22.8%

(a) Males

Logistic regression was used to assess the strength of association between sociodemographic, HIV awareness and sexual behaviour characteristics and risk of prevalent HIV infection. Among males, since there were no HIV infections among one village all risk factor analyses had to adjust only for village pair. HIV prevalence was significantly associated with age after adjustment for village pair (Table 7.7, aOR 2.61 95% CI 1.40-4.85). HIV prevalence was higher in non-Sepedi speaking individuals (10.1% vs. 5.2%), but this association did not persist after adjustment for village and other confounders. There was no association between prevalent HIV infection and place of birth or travel to a city during the past year among males. Males who came from households with more than two other young people aged 14-25 years were less likely to be infected with HIV but this association did not reach statistical significance.

HIV serostatus was not associated with knowledge about HIV or communication. There were no infections among the 32 individuals who had had an HIV test and it was thus not possible to conduct further statistical analysis among this group. Those who held more stigmatising attitudes towards those living with HIV were less likely to be HIV infected after adjustment for confounding factors (aOR 0.52 95% CI 0.27-1.00).

Some 3.2% of males who reported never having had sexual intercourse were HIV positive. This issue is discussed in greater detail in Chapter 9. Individuals reporting some sexual partners were more likely to be infected with HIV than those who reported none. This association was attenuated by the introduction of the age and village variables into the logistic regression model. Young age and not wanting to have sex at sexual debut were not associated with prevalent HIV infection. Individuals who were practising secondary abstinence were more likely to be HIV infected, and this association was of borderline significance after adjustment for confounding variables (aOR 0.39 95% CI 0.14-1.06).

Circumcision was common among the male population and was not associated with HIV infection. Among those individuals who had been sexually active during the past year, those reporting any partnerships where sex was infrequent were less likely to be HIV positive (aOR 0.39 95% CI 0.15-1.04). Other characteristics of partnerships in the last year including age difference with partners, resource exchange and condom use were not associated with HIV prevalence after adjustment for confounding variables.

Table 7.7: HIV risk factor analysis among males

Risk Factor	Group	HIV positive N=763			aOR 1	aOR 2	aOR 3 (95% CI)		aOR 3 Adjusted for age, village and ...
Selected sociodemographic factors		n	N	%					
Age	14-19	19	504	3.8			1		-
	20-25	24	259	9.3			2.61	1.40-4.85	
Language	Sepedi	36	694	5.2	1	1	1		Number of 14-25 year olds in home
	Other	7	69	10.1	2.03	1.87	1.78	0.72-4.40	
Orphanhood	Both parents alive	26	507	5.1	1	1	1		
	One or both parents dead or missing	17	256	6.6	1.15	1.26	1.24	0.63-2.42	
Born outside village of current residence	No	29	581	5.0	1	1	1		
	Yes	14	182	7.7	1.47	1.06	1.07	0.51-2.25	
Number of 14-25 year olds in home	1-2	21	325	6.5	1	1	1		
	3+	22	438	5.0	0.66	0.59	0.59	0.31-1.13	
Travel to a city in past year	No	30	529	5.7	1	1	1		
	Yes	13	234	5.6	0.72	0.83	0.85	0.42-1.73	
HIV awareness									
HIV Knowledge	High	23	328	7.0	1	1	1		Number of 14-25 year olds in home, Stigmatising attitudes
	Low	20	435	4.6	0.77	0.94	1.04	0.54-2.00	
Household Communication	Yes	19	305	6.2	1	1	1		
	No	24	458	5.2	1.10	1.12	1.13	0.58-2.21	
HIV tested	Yes	0	32	0.0	-	-	-		
	No	43	731	5.9	-	-	-		
Stigmatising Attitudes	Tolerant	25	299	8.4	1	1	1		
	Stigmatising	18	464	3.9	0.54	0.53	0.52	0.27-1.00	
Sexual behaviour factors									
Age of sexual debut+	Under 15 yrs	33	618	5.3	1	1	1		Number of 14-25 year olds in home, Stigmatising attitudes, Number of lifetime partners, Number of partners in last year
	Over 14 yrs	10	133	7.5	1.57	1.82	1.81	0.76-4.32	
Experience at sexual debut*	Wanted to have sex	28	378	7.4	1	1	1		
	Didn't want to have sex	6	99	6.1	0.97	1.03	1.19	0.45-3.15	
No of lifetime partners	None	9	286	3.2	1	1	1		
	1-3	17	243	7.0	1.47	1.32	2.41	0.74-7.84	
	4 or more	17	234	7.3	1.14	0.99	2.41	0.62-9.39	
No of partners in last year	0	16	329	4.9	1	1	1		
	1	21	300	7.0	0.75	0.77	0.47	0.17-1.29	
	More than one	6	134	4.5	0.43	0.45	0.28	0.08-1.03	
Secondary Abstinence*	Yes	7	43	16.3	1	1	1		
	No	27	434	6.2	0.26	0.35	0.39	0.14-1.06	
Co-factor for transmission									
Circumcised	No	39	666	5.9	1	1	1		
	Yes	4	97	4.1	0.80	0.94	1.38	0.37-5.21	
Partnership characteristics*									
Frequency of intercourse	All partners sex more than 5 times in past year	20	215	9.3	1	1	1		Number of 14-25 year olds in home, Stigmatising attitudes, Number of lifetime partners, Number of partners in last year, Frequency of intercourse
	Any partners <5times sex in past year	7	219	3.3	0.38	0.37	0.39	0.15-1.04	
Resource exchange between partners in last year	All partnerships no, reciprocal or female to male resource exchange	20	296	6.8	1	1	1		
	At least one partnership male provides resources but female does not	7	138	5.1	0.66	0.61	0.55	0.21-1.42	
Age difference with partners in last year	All more than 3 years younger	10	103	9.7	1	1	1		
	Any of similar age or older	16	315	5.1	0.77	0.86	0.89	0.34-2.31	
Condom use in last year	Ever	11	196	5.6	1	1	1		
	Never	16	238	6.7	1.31	1.24	1.82	0.75-4.40	

aOR_1= Odds Ratio (95% Confidence Intervals) adjusted for age, aOR_2 . adjusted for age and village pair, aOR_3 Adjusted for age, village pair and other factors retained in stepwise models run at each hierarchical stage.
* Only among those sexually active during the previous 12 months

(b) Females

Among females, HIV infection was also significantly associated with age (Table 7.8). Generally, HIV infection was not associated with sociodemographic factors. However, females who had travelled to a city were at increased risk of infection (aOR 1.38 95% CI 0.68-2.21). Individuals co-resident with two or more other 14-25 year olds had lower levels of prevalent HIV infection with borderline statistical significance (aOR 0.71 95% CI 0.46-1.09), a finding also seen among the males.

Of the HIV awareness variables, HIV knowledge and communication about sex/sexuality were not associated with HIV prevalence. Those who had had a test for HIV were more likely to be infected although the numbers were small. As for the males, holding stigmatising attitudes towards people living with HIV/AIDS was associated with a significant reduction in risk of prevalent HIV infection (aOR 0.61 95% CI 0.40-0.93).

As for the males, there were some reported HIV infections among females who reported never having had sexual intercourse (15/263, 5.7%). Having had a larger number of lifetime sexual partners was significantly associated with risk of infection after adjustment for age and other confounders (aOR 2.55, 95% CI 1.13-5.76). However, other characteristics were not associated with HIV prevalence.

Table 7.8: HIV risk factor analysis among females

Risk Factor	Group	HIV positive N=763			aOR ₁	aOR ₂	aOR (95% CI) ₃		aOR ₃ Adjusted for age, village and ...
Selected sociodemographic factors		n	N	%					
Age	14-19	44	557	7.9			1		Number of 14-25 year olds in home, Travel to a city in past year
	20-25	64	323	19.8			2.88	1.91-4.35	
Language	Sepedi	93	795	11.7	1	1	1		
	Other	15	85	17.7	1.56	1.41	1.45	0.76-2.77	
Orphanhood	Both parents alive	68	608	11.2	1	1	1		
	One or both parents dead or missing	40	272	14.7	1.17	1.27	1.27	0.82-1.97	
Born outside village of current residence	No	76	641	11.9	1	1	1		
	Yes	32	239	13.4	1.10	0.96	0.92	0.57-1.49	
Number of 14-25 year olds in home	1-2	47	348	13.5	1	1	1		
	3+	61	532	11.5	0.72	0.70	0.71	0.46-1.09	
Travel to a city in past year	No	74	663	11.2	1	1	1		
	Yes	34	217	15.7	1.30	1.40	1.38	0.86-2.21	
HIV awareness									
HIV Knowledge	High	56	387	14.5	1	1	1		Number of 14-25 year olds in home, Travel to a city in past year, Stigmatising attitudes
	Low	52	493	10.6	0.80	0.82	0.91	0.59-1.40	
Household Communication	Yes	59	426	13.9	1	1	1		
	No	49	454	10.8	0.83	0.88	0.93	0.60-1.44	
HIV tested	Yes	17	78	21.8	1	1	1		
	No	91	802	11.4	0.61	0.58	0.60	0.32-1.12	
Attitudes	Tolerant	55	337	16.3	1	1	1		
	Stigmatising	53	543	9.8	0.60	0.60	0.61	0.40-0.93	
Sexual behaviour factors									
Age of sexual debut	Under 15 yrs	95	790	12.0	1	1	1		Number of 14-25 year olds in home, Travel to a city in past year, Stigmatising attitudes, Number of lifetime partners
	Over 14 yrs	12	86	14.0	1.32	1.21	1.14	0.57-2.27	
Experience at sexual debut*	Wanted to have sex	52	313	16.6	1	1	1		
	Didn't want to have sex	41	304	13.5	0.84	0.82	0.83	0.52-1.34	
No of lifetime partners	None	15	263	5.7	1	1	1		
	1-3	42	352	11.9	1.94	1.91	1.81	0.86-3.80	
	4 or more	51	265	19.3	2.81	2.71	2.55	1.13-5.76	
No of partners in last year	0	26	325	8.0	1	1	1		
	1	75	503	14.9	1.35	1.49	1.30	0.72-2.35	
	More than one	7	52	13.5	1.36	1.61	1.31	0.48-3.55	
Had any children	No	54	562	9.6	1	1	1		
	Yes	54	318	17.0	1.04	1.02	0.94	0.55-1.59	
Secondary Abstinence*	Yes	11	62	17.7	1	1	1		
	No	82	555	14.8	0.78	0.97	0.87	0.41-1.86	
Among sexually active in the past year									
Frequency of intercourse	All partners sex more than 5 times in past year	55	351	15.7	1	1	1		Number of 14-25 year olds in home, Travel to a city in past year, Stigmatising attitudes, Number of lifetime partners, Resource exchange between partners in the last years, Age difference with partners in the last year
	Any partners <5times sex in past year	27	204	13.2	1.05	1.09	1.17	0.65-2.08	
Resource exchange between partners in last year	All partnerships no, reciprocal or female to male resource exchange	15	144	10.4	1	1	1		
	At least one partnership male provides resources but female does not	67	411	16.3	1.86	1.84	1.67	0.86-3.23	
Age difference with partners in last year	All > 3 years younger	26	236	11.0	1	1	1		
	Any similar age / older	50	289	17.3	1.51	1.54	1.42	0.83-2.43	
Condom use in last year	Ever	38	282	13.5	1	1	1		
	Never	44	273	16.1	1.25	1.25	1.37	0.80-2.32	

aOR_1= Odds Ratio (95% Confidence Intervals) adjusted for age, aOR_2 - adjusted for age and village, aOR_3 Adjusted for age, village and other factors retained in stepwise models run at each hierarchical stage.

Among the group reporting sexual partnerships during the previous year, there was no association between the condom use variable and risk of prevalent infection. Individuals reporting any partnerships in the last year characterised by a wide age gap or male to female only resource exchange were at higher risk of HIV infection though this was not statistically significant at $p<0.05$.

7.04 Summary of key findings

Lack of basic knowledge about HIV and lack of communication with parents or other household members were widespread. HIV testing was rare, with some of those who had been tested not requesting a test and not receiving the results. A high number of individuals held some stigmatising views on those living with HIV/AIDS.

More than half of the individuals interviewed had previously been sexually active, with this being strongly associated with age. Secondary abstinence was relatively rarely reported. Above 17 years of age about one quarter of males reported multiple partners during the previous year, while among females both lifetime and recent numbers of reported sexual partners were lower than for males. Partnerships reported by young people commonly featured significant male seniority in terms of age and some resource exchange between partners, which was more often conceptualised as being from male-to-female partner by females. Condom use within partnerships was not strongly influenced by age of the reporting individual, but was influenced by the frequency of sexual intercourse within partnerships among both sexes and by resource exchange patterns among relationships reported by females.

Prevalent HIV infection was significantly more common among females than males and was strongly associated with increased age among both sexes. Relatively few measured sociodemographic, HIV awareness or sexual behaviour characteristics were associated with prevalent HIV infection among either sex. However, holding stigmatising attitudes was associated with a lower prevalence of HIV infection among both sexes. Among females, but not males, increased numbers of reported lifetime partners was associated with risk of infection after adjustment for age.

Chapter 8. The social epidemiology of sexual behaviour and HIV infection

8.01 Introduction

The analyses described in the previous chapters provided a good understanding of the distribution of both exposures and outcomes in the study population. The remaining stage of the thesis was to examine the association between these factors, following the statistical analysis strategy described in Chapter 4.

8.02 Objectives of chapter

1) Describe the association between household wealth and HIV awareness, sexual behaviour and HIV infection among males and females

2) Describe the association between socioeconomic role variables and HIV awareness, sexual behaviour and HIV infection among males and females

3) Assess whether there was interaction in the relationship between household wealth and socioeconomic roles in their association with outcome variables, or whether differences in the distribution of measured pathway variables were responsible for mediating the relationships seen between household wealth and socioeconomic roles, HIV awareness, sexual behaviour and HIV infection.

8.03 Results

8.03.1 Household wealth and HIV risk characteristics

(a) Males

Household wealth was not generally associated with HIV risk characteristics among males. Only 2/16 of the odds ratios presented comparing those from "poor, but a bit better off" households compared to those from the "very poor" households were less than 1, while 9/16 of those comparing males from households that were "doing OK" with the "very poor" group were less than 1. The confidence intervals around only two effect estimates did not include 1 since effect sizes were generally modest. One of these odds ratios and 95% confidence intervals suggested higher risk characteristics among the middle "poor, but a bit better off" compared to the very poor group (Table 8.1).

Table 8.1: The association between household wealth and HIV risk characteristics among males aged 14-25 years

Variable	Proportion with high risk category												
	Very poor		Poor, but a bit better off					Doing OK					p
	n	%	n	%	aOR1	aOR2 (95%CI)		n	%	aOR1	aOR2 (95% CI		(Age interaction)
N	272		524					120					
HIV awareness													
Lack of HIV knowledge	163	59.9	305	58.2	0.94	1.03	0.75-1.41	61	50.8	0.66	0.74	0.47-1.18	0.035
14-25 years	102/170	60.0	218/332	65.7	1.32	1.46	0.98-2.19	44/75	58.7	0.89	1.03	0.57-1.83	-
20-24 years	61/102	59.8	87/192	45.3	0.56	0.57	0.34-0.96	17/45	37.8	0.41	0.44	0.20-0.96	-
Lack of household communication	172	63.2	309	59.0	0.84	0.77	0.56-1.06	67	55.8	0.67	0.67	0.42-1.08	0.316
Not had an HIV test	251	92.3	505	96.4	2.28^	1.75^	0.88-3.46	112	93.3	1.11^	1.03^	0.41-2.58	0.139
Holds stigmatising attitudes	165	60.7	317	60.5	1.02	1.04	0.76-1.44	59	49.2	0.57	0.59	0.37-0.95	0.976
Individual level sexual behaviour													
Ever had sex *	165	60.7	337	64.3	0.98	0.93	0.77-1.14	76	63.3	0.99	0.89	0.67-1.20	0.575
Unwanted sex at debut $	34	20.6	67	19.9	0.92	1.06	0.65-1.74	16	21.1	0.98	0.91	0.44-1.88	0.531
More than 3 partners during lifetime	84	30.9	173	33.0	1.12	1.09	0.75-1.56	37	30.8	1.10	1.06	0.61-1.84	0.952
More than 1 partner during previous year	47	17.3	99	18.9	1.09	1.13	0.75-1.70	18	15.0	0.90	0.86	0.46-1.63	0.611
Continuing sexual activity (opp. 2° abstinence) $	149	90.3	304	90.2	1.01	1.21	0.62-2.37	70	92.1	1.41	1.72	0.60-4.93	0.792
Co-factor													
Uncircumcised	27	9.9	63	12.0	1.28	1.16	0.70-1.92	18	15.0	1.53	1.44	0.73-2.85	0.566
Partnership characteristics *(N)*	212		443					93					
Intercourse > 5 times in past year	128	60.4	263	59.4	0.99	1.00	0.70-1.45	47	50.5	0.72	0.66	0.39-1.14	0.482
Partners of similar ages	127	65.5	291	69.1	1.15	1.19	0.75-1.89	63	72.4	1.29	1.39	0.68-2.85	0.253
Only male partner provides resources	56	26.4	130	29.4	1.15	1.07	0.73-1.58	21	22.6	0.80	0.71	0.39-1.32	0.438
Never used a condom in the last year	91	42.9	223	50.3	1.38	1.47	1.03-2.09	41	44.1	0.96	1.03	0.61-1.78	0.036
14-25 years	48/92	52.2	101/203	49.8	0.95	1.05	0.55-1.98	22/40	55.0	0.92	0.70	0.27-1.82	-
20-24 years	43/120	35.8	122/240	50.8	1.86	1.81	1.01-3.26	19/53	35.9	1.00	1.05	0.44-2.49	-
No condom use at last sex	140	66.0	305	68.9	1.15	1.19	0.81-1.72	56	60.2	0.74	0.71	0.41-1.22	0.001
14-25 years	71/92	77.2	133/203	65.5	0.57	0.49	0.27-0.92	28/40	70.0	0.63	0.48	0.19-1.19	-
20-24 years	69/120	57.5	172/240	71.7	1.88	2.42	1.44-4.05	28/53	52.8	0.83	1.12	0.54-2.34	-
HIV *(N)*	227		431					105					
HIV Positive	10	4.4	26	6.0	1.38^	1.25^	0.57-2.73	7	6.7	1.64^	1.44^	0.50-4.12	0.462

aOR1 adjusted for age, aOR2 adjusted for age, village, number of co-resident young people, current residency in village of birth, orphanhood. Missing data was on age of sexual debut (15) and age difference between partners (42). * Effect estimate was Hazard ratio from Cox regression. $ Only relevant to individuals who had previously ever had sex (see table for numbers). ^ Adjustment for village pair only. Underlined when likelihood ratio test p<0.05, or when 0.05<p<0.10.

There was some evidence that those from the wealthiest group had greater HIV-related knowledge although this association was not statistically significant. This relationship also showed a significant age interaction. Among the younger group, those in the middle wealth band more often reported a lack of HIV knowledge than those in the "very poor" or "doing OK" group, while among the older group both of the wealthier groups were less likely to report a lack of knowledge. The variable "lack of household communication" showed some evidence of a trend across groups, with those in the middle and "doing OK" groups reporting this less often than those from "very poor" households with borderline significance in each case. Those from households that were "doing OK" were significantly less likely to report holding stigmatising attitudes than the "very poor".

None of the individual level sexual behaviour variables was associated with household wealth. Similarly, circumcision status did not vary across wealth groups.

Among characteristics measured for sexual partnerships in the last year, there was some weak evidence that those from the "doing OK" group were less likely to report partnerships with more frequent sexual activity than the "very poor" group. Never having used a condom within partnerships was more often reported by those in the middle wealth group compared to the poorest. There was some evidence of age interaction. No association was noted in the younger group, while among the older group lack of condom use was most common among those in the middle wealth band. Conversely, lack of condom use at last sex was least often reported by those from the wealthiest households although this was not statistically significant. There was also evidence of age interaction in this relationship. Among those in the younger age group lack of condom use was least common among the two wealthier groups. Among the older age group, not having used a condom at last sex was more often reported by the middle than the "very poor" group, but there was little difference between the very poor and wealthiest groups. Finally, HIV infection, which was quite rare among young males, also showed little evidence for variation across wealth groups.

The age-stratified models for each of the condom use variables were further adjusted for the stigmatising attitudes variable to investigate mediation. When never having used a condom in the past year was the outcome variable this additional adjustment had no effect on the odds ratios associated with the two wealthier groups among those aged 14-19 years. Among the older group the odds ratios did shift slightly away from one, such that for the "poor, but a bit better off" group the new odds ratio comparing to the "very poor" was 1.94 (95% CI 1.07-3.53), while for the "doing OK" group the new odds ratio was 1.18 (95% CI 0.49-2.83). When not having used a condom at last sex was the outcome variable there was no change in the odds ratios in either age strata.

(b) Females

Comparing the "poor, but a bit better off" group with the "very poor" group, 9/16 odds ratios suggested lower risk in the middle group, while 11/16 point estimates suggested lower risk in the group from households that were "doing OK". However, most effects were small and not statistically significant.

Table 8.2: The association between household wealth and HIV risk characteristics among females aged 14-25 years

Variable	Proportion with high risk category												
	Very poor		Poor, but a bit better off					Doing OK					p
	n	%	N	%	aOR1	aOR2 (95%CI)		n	%	aOR1	aOR2 (95% CI		(Age interaction)
N	**296**		**548**					**159**					
HIV awareness													
Lack of HIV knowledge	168	56.8	302	55.1	0.90	**0.87**	**0.65-1.17**	91	57.2	0.98	**0.98**	**0.65-1.48**	0.787
Lack of household communication	161	54.4	276	50.4	0.83	**0.75**	**0.56-1.01**	84	52.8	0.92	**0.87**	**0.58-1.31**	0.549
Not had an HIV test	271	91.6	492	89.8	0.72	**0.66**	**0.38-1.12**	146	91.8	0.95	**0.93**	**0.44-1.96**	0.108
Holds stigmatising attitudes	185	62.5	349	63.4	1.02	**1.03**	**0.76-1.39**	86	54.1	0.67	**0.72**	**0.48-1.08**	0.340
Individual level sexual behaviour													
Ever had sex *	214	72.3	384	70.1	1.05	**1.05**	**0.88-1.25**	108	67.9	0.96	**0.95**	**0.74-1.21**	0.498
Unwanted sex at debut $	101	47.2	190	49.5	1.03	**1.16**	**0.81-1.66**	55	50.9	1.14	**1.36**	**0.83-2.24**	0.211
More than 2 partners during lifetime	93	31.4	166	30.3	1.15	**1.19**	**0.83-1.71**	48	30.2	1.13	**1.21**	**0.74-1.99**	0.766
More than 1 partner during previous year	14	4.7	36	6.6	1.43	**1.62**	**0.84-3.10**	7	4.4	0.94	**1.10**	**0.42-2.87**	0.619
Had a child	124	41.9	183	33.4	0.78	**0.76**	**0.52-1.10**	55	34.6	0.82	**0.73**	**0.43-1.23**	0.772
Continuing sexual activity (opp. 2° abstinence) $	195	91.1	347	90.4	0.93	**0.92**	**0.50-1.71**	94	87.0	0.68	**0.83**	**0.38-1.84**	0.876
Partnership characteristics *(N)*	**211**		**387**					**102**					
Intercourse > 5 times in past year	134	63.5	259	66.9	1.29	**1.40**	**0.96-2.05**	71	69.6	1.35	**1.45**	**0.85-2.50**	0.971
Male partner more than 3 yrs older than female	105	55.3	192	51.9	0.93	**0.94**	**0.65-1.36**	58	59.8	1.25	**1.26**	**0.75-2.11**	0.530
Only male partner provides resources	157	74.4	271	70.0	0.76	**0.84**	**0.56-1.25**	72	70.6	0.78	**0.85**	**0.49-1.49**	0.985
Never used a condom in the last year	110	52.1	186	48.1	0.84	**0.82**	**0.57-1.16**	43	42.2	0.65	**0.64**	**0.39-1.06**	0.193
No condom use at last sex	157	74.4	276	71.3	0.89	**0.84**	**0.56-1.25**	61	59.8	0.50	**0.47**	**0.28-0.80**	0.146
HIV *(N)*	**254**		**482**					**144**					
HIV Positive	32	12.6	61	12.7	1.14	1.14	0.71-1.15	15	10.4	0.91	**0.88**	**0.44-1.75**	0.507

aOR1 adjusted for age, aOR2 adjusted for age, village, number of co-resident young people, current residency in village of birth, orphanhood. Missing data was on age of sexual debut (7) and age difference between partners (45). * Effect estimate was Hazard ratio from Cox regression. $ Only relevant to individuals who had previously ever had sex (see table for numbers). Underlined when likelihood ratio test p<0.05, or when 0.05<p<0.10.

HIV knowledge was not significantly differently distributed across wealth groups. The variable "lack of communication" was less common among the middle group compared to the very poor. Individuals in the "poor, but a bit better off" group were also less likely not to have had an HIV test. Holding stigmatising attitudes was reported less often by those from households that were "doing OK". In each case, however, likelihood ratio tests comparing models with and without the wealth variable in three groups gave significance values of $p>0.10$.

None of the individual level sexual behaviour characteristics were associated with household wealth. Among sexual partnerships reported for the last year by those who had been sexually active, frequency of intercourse, age difference and resource exchange also showed no variation across wealth groups. However, both condom use variables suggested lower risk among females from households that were "doing OK" compared to both the poorest and middle groups. There was no association between household wealth and either having had a child or HIV prevalence.

Finally, there were no age interactions for the relationships between wealth and any outcome variables that were significant at $p<0.05$ among females. Since only one outcome variable (condom use at last sex) was associated with household wealth, it was not necessary to examine whether variables higher up the causal pathway mediated this association.

8.03.2 Socioeconomic Roles and HIV risk characteristics

(a) Males

1) School attendance

Some 12/16 of the associations suggested lower risk characteristics among those attending school. Generally effect sizes were modest. Four of these associations were statistically significant.

All of the HIV awareness variables examined suggested lower risk characteristics among those in school, but none were statistically significant. Only one factor suggested an age interaction with borderline significance. Among the younger age group, there was a significantly lower prevalence of individuals holding stigmatising attitudes among those in school (282/506, 55.7% vs. 41/61, 67.2%, aOR 0.36 95% CI 0.36-0.69). Among the older group there was some weak evidence to suggest a higher prevalence of this stigmatising attitude among those in school (46/110, 41.8% vs. 71/224 31.7%, aOR 1.41 95% CI 0.80-2.48).

Among individual level sexual behaviour characteristics, having had a high number of partners during the lifetime or in the past year were both significantly less commonly reported by those in school. Ever having had sexual intercourse, type of sexual debut and secondary abstinence among those previously sexually active were not significantly associated with current attendance at school. Similarly, circumcision status was not associated with attendance at school.

Table 8.3: School attendance and HIV awareness, sexual behaviour and HIV infection among men aged 14-25

Variable	Not attending school		Attending school		aOR1	aOR2		aOR3	P (Age Interaction)
						aOR	**95% CI**		
	n	%	n	%					
N	**288**		**628**						
HIV awareness									
Lack of HIV knowledge	154	53.5	375	59.7	0.77	**0.83**	**0.57-1.21**	0.80	0.225
Lack of household communication	149	51.7	399	63.5	0.79	**0.75**	**0.51-1.10**	0.70	0.436
Not had an HIV test	266	92.4	602	95.9	0.65	**0.59^**	**0.28-1.25**	0.58^	0.157
Holds stigmatising attitudes	144	50.0	397	63.2	0.87	**0.88**	**0.60-1.28**	0.79	0.072
Individual level sexual behaviour									
Ever had sex *	250	86.8	328	52.2	0.96	**0.93**	**0.75-1.15**	0.92	0.997
Unwanted sex at debut $	41	16.4	76	23.2	1.03	**0.99**	**0.58-1.71**	1.00	0.221
More than 3 partners during lifetime	158	54.9	136	21.7	0.67	**0.66**	**0.44-0.97**	0.67	0.210
More than 1 partner during previous year	81	28.1	83	13.2	0.63	**0.61**	**0.39-0.94**	0.60	0.936
Continuing sexual activity (opp. 2° abstinence) $	229	91.6	294	89.6	1.14	**1.22**	**0.57-2.61**	1.29	0.625
Co-factor									
Uncircumcised	20	6.9	88	14.0	1.95	**1.79**	**0.89-3.57**	1.77	0.695
Partnership characteristics *(N)*	345		403						
Intercourse > 5 times in past year	232	67.3	206	51.1	0.75	**0.75**	**0.51-1.10**	0.81	0.404
Partners of similar ages	153	50.9	318	83.3	1.13	**1.09**	**0.67-1.69**	1.10	0.154
Only male partner provides resources	108	31.3	99	24.6	0.68	**0.65**	**0.43-0.99**	0.68	0.802
Never used a condom in the last year	157	45.5	198	49.1	0.94	**1.00**	**0.69-1.47**	0.96	0.322
No condom use at last sex	232	67.3	269	66.8	0.80	**0.89**	**0.59-1.32**	0.84	0.274
HIV *(N)*	213		550						
HIV Positive	25	11.7	18	3.3	0.35^	**0.27^**	**0.11-0.65**	0.27^	0.138

aOR1 adjusted for age, aOR2 adjusted for age, confounders and household wealth, aOR3 adjusted for age, confounders, wealth and temporary migrancy. Missing data was on age of sexual debut (15) and age difference between partners (42). * Effect estimate was Hazard ratio from Cox regression. ^ Adjustment for village pair only. $ Only among those who had previously had sex. Underlined when likelihood ratio test $p<0.05$, or when $0.05<p<0.10$.

Partnerships during the last year characterised by male to female only resource exchange were significantly less commonly reported by males attending school / college. There was some weak evidence that partnerships where sexual intercourse had occurred more than five times were relatively less commonly reported by those in school, although in a likelihood ratio test this did not reach statistical significance at the level of $p<0.05$.

There was strong evidence among males to suggest that current school attendance was statistically significantly associated with a lower prevalence of HIV infection.

Further adjustment for temporary migrancy status in each of the models presented in Table 8.3 did not suggest a major confounding influence of migrancy on the associations presented.

The sexual behaviour terms significantly associated with school attendance were included into the model in which HIV was the outcome. Including terms for numbers of lifetime and recent partners had almost no effect on the odds ratio associated with schooling (aOR 0.26 95% CI 0.11-0.64). If the analysis was restricted only to individuals who had had a sexual partner during the past year, the association between schooling and HIV infection was no longer significant (aOR 0.44 95% CI 0.15-1.32), although the odds ratio was still lower than one, but with wider confidence intervals due to the lower number of individuals included in the analysis. Adjusting the model for resource exchange patterns had a limited impact on the odds ratio describing the association between school attendance and HIV infection (aOR 0.40 95%CI

0.12-1.23), suggesting that this factor did not mediate the difference in HIV prevalence that was noted between those in and out of school.

2) Temporary migrancy

Some 9/16 odds ratios suggested lower risk characteristics among temporary migrants than among those who were non-migrant. The confidence intervals for three odds ratios did not include 1, with two of these suggesting lower risk characteristics in the temporary migrant group.

Temporary migrants were less likely to have low HIV knowledge than non-migrants. Temporary migrants were also significantly less likely not to communicate with their parents / guardians or other household members on issues such as sex and sexuality or to hold stigmatizing attitudes towards HIV positive individuals.

Individual sexual behaviour characteristics and circumcision were not associated with temporary migrancy, although temporary migrants were slightly less likely to report that they did not want to have sex at sexual debut. There was a borderline significant age interaction between migrancy and age for the circumcision variable, although numbers of uncircumcised individuals were very low (2) in one cell rendering this test unstable.

Table 8.4: Temporary migration, HIV awareness, sexual behaviour and HIV infection among men aged 14-25

Variable	Proportion with high risk				aOR1	aOR2		aOR3	P (Age Interaction)
	Non-migrant		Migrant						
	n	%	n	%					
N	**788**		**128**						
HIV awareness									
Lack of HIV knowledge	468	56.4	61	47.7	0.78	**0.74**	**0.49-1.11**	0.71	0.428
Lack of household communication	492	62.4	56	43.8	0.66	**0.64**	**0.42-0.96**	0.61	0.540
Not had an HIV test	752	95.4	116	90.6	0.76^	**0.81^**	**0.39-1.70**	0.77^	0.318
Holds stigmatising attitudes	492	62.4	49	38.2	0.51	**0.47**	**0.31-0.71**	0.45	0.676
Individual level sexual behaviour									
Ever had sex *	473	60.0	105	82.0	1.00	**0.99**	**0.78-1.24**	0.98	0.790
Unwanted sex at debut $	103	21.8	14	13.3	0.68	**0.59**	**0.31-1.13**	0.58	0.835
More than 3 partners during lifetime	232	29.4	62	48.4	1.12	**1.22**	**0.79-1.88**	1.15	0.902
More than 1 partner during previous year	133	16.9	31	24.2	1.07	**1.04**	**0.64-1.69**	0.97	0.919
Continuing sexual activity (opp. 2° abstinence) $	424	89.6	99	94.3	1.64	**1.52**	**0.60-3.87**	1.59	0.473
Co-factor									
Uncircumcised	97	12.3	11	8.6	0.85	**0.84**	**0.42-1.71**	0.89	0.055
Partnership characteristics *(N)*	**609**		**139**						
Intercourse > 5 times in past year	336	55.2	102	73.6	1.79	**1.85**	**1.19-2.86**	1.78	0.989
Partners of similar ages	402	70.8	79	58.9	1.35	**1.23**	**0.77-1.97**	1.25	0.675
Only male partner provides resources	160	26.3	47	33.8	1.41	**1.44**	**0.94-2.20**	1.35	0.591
Never used a condom in the last year	299	49.1	56	40.3	0.76	**0.77**	**0.52-1.15**	0.77	0.829
No condom use at last sex	416	68.3	85	61.2	0.76	**0.76**	**0.50-1.14**	0.74	0.450
HIV *(N)*	**664**		**99**						
HIV Positive	34	5.1	9	9.1	1.23^	**1.36^**	**0.58-3.19**	1.17^	0.019
14-19years	15/470	3.2	4/34	11.8	3.55+	3.98+	1.18-13.40	2.61+	-
20-25 years	19/194	9.8	5/65	7.7	0.70	0.70	0.24-2.05	0.67	-

aOR1 adjusted for age, aOR2 adjusted for age, confounders and household wealth, aOR3 adjusted for age, confounders, wealth and attendance at school. Missing data was on age of sexual debut (15) and age difference between partners (42). * Effect estimate was Hazard ratio from Cox regression. ^Adjustment for village pair only. +No adjustment for village or pair $Only among those who had previously had sex. Underlined when likelihood ratio test p<0.05, or 0.05<p<0.10.

Sexual partnerships reported by temporary migrants were significantly more likely to be partnerships where sex had occurred frequently during the previous year. Migrants were also more likely to report relationships characterized by one way male-to-female resource flow as opposed to any other resource exchange pattern, with this relationship being of borderline statistical significance. There was little evidence of an association between migrancy and age difference within sexual partnerships. Both "lack of condom use" variables were more often reported by non-migrants, although neither was significantly associated with this.

Another statistically significant age interaction was seen in the relationship between migrancy and HIV serostatus (p=0.019). Among the whole group there was little evidence of a relationship between temporary migrancy and HIV infection. After stratification by age there were smaller numbers in each strata. There was no association between temporary migrancy and HIV serostatus among the older group. Among the younger group there was a significantly higher risk of HIV infection among temporary migrants. However, this estimate was based on small numbers of individuals (only 4 migrants were HIV positive), and it was not possible to adjust this estimate for village or pair due to low numbers.

Adjusting the association between temporary migrancy and intercourse frequency for the HIV awareness variables significantly associated with migrancy status had very little impact on the odds ratio (aOR 1.88 95% CI 1.21-2.92), suggesting that these pathway variables did not mediate the association seen.

Finally among the males, there was little evidence for interaction between school attendance and temporary migrancy in their association with most HIV risk characteristics (Table 8.5). There was some suggestion of an interaction between schooling and migrancy in their relationship with ever having had an HIV test, although few individuals had ever had a test. Among those who were non-migrants there was no association between schooling and HIV testing (of those out of school 189/207, 91.3%, had not had an HIV test compared to 563/581, 96.9% of those in school, aOR 1.17 95% CI 0.50-2.73). However, among the migrants, never having had an HIV test was significantly less common among those attending school (77/81, 95.1% of those out of school vs. 39/47, 83.0% of those in school, aOR 0.11 95% CI 0.02-0.55). There was also evidence of an interaction between migrancy and schooling for the HIV infection outcome. Among the majority non-migrants, reduced risk of HIV infection was strongly associated with attendance at school (of those out of school 21/152, 13.8%, were HIV positive compared to 13/512, 2.5% of those in school, aOR 0.23 95% CI 0.09-0.60), while among the smaller group of migrants there was little evidence of an association, though this was based on small numbers in some cells (4/61, 6.6% of those out of school vs. 5/38, 13.2% of those in school, aOR 1.90 95% CI 0.34-10.46).

Table 8.5: Likelihood ratio tests for interaction between migration and school attendance in their association with HIV risk outcomes

Variable	Males	Females
	p	p
HIV awareness		
Lack of HIV knowledge	0.954	0.824
Lack of household communication	0.403	0.326
Not had an HIV test	<u>0.012</u>	0.474
Holds stigmatising attitudes	0.189	0.267
Individual level sexual behaviour		
Ever had sex *	0.346	0.481
Unwanted sex at debut $	0.735	<u>0.017</u>
More than 3 (male) or 2 (female) partners during lifetime	0.358	0.072
More than 1 partner during previous year	0.280	0.990
Childbirth	-	0.055
Continuing sexual activity (opp. 2° abstinence) $	0.162	<u>0.038</u>
Co-factor		
Uncircumcised	0.204	-
Partnership characteristics *(N)*		
Intercourse > 5 times in past year	0.402	0.853
Partners of similar ages (males) Male partner more than 3 yrs older than female (females)	0.263	0.892
Only male partner provides resources	0.310	0.843
Never used a condom in the last year	0.378	0.332
No condom use at last sex	0.135	0.076
HIV *(N)*		
HIV Positive	<u>0.001</u>	0.640

Likelihood ratio tests comparing models adjusted for all confounding factors with and without interaction term. * Cox regression. $ Only among those who had previously had sex. Underlined when likelihood ratio test <u>p<0.05</u>, or when 0.05<p<0.10.

To check whether the interaction between migrancy and schooling was partly caused by the association between migrancy and age, tests were also conducted for this interaction in each of the age groups separately. When this was done there were very small numbers in some cells and the power of these tests were very low. Nevertheless there was some evidence that the interaction between migrancy and schooling was seen in both the younger (p=0.013) and older (p=0.09) groups.

(b) Females

1) School attendance

Among females, overall more point estimate odds ratios suggested lower (13/16) than higher (3/16) risk characteristics among those attending school compared with those not doing so after adjustment for childbirth. Effect sizes were moderate as for the males, and 95% confidence intervals around the odds ratios for four reported relationships excluded 1.

None of the HIV awareness variables was significantly associated with school/college attendance, although lack of household communication was more common among those in school with borderline significance before adjustment for childbirth. One variable, having had an HIV test, showed borderline significant age interaction. Among both age groups however there was only weak evidence of an association. Among the younger group, 78/90, 86.7% of those out of school had not had a test vs. 510/531, 96.1% of those in school (aOR 1.15 95% CI 0.43-3.07). Among the older group, 246/290, 84.8% had not tested among those out of school vs. 75/92, 81.5%, of those in school (aOR 0.64 95% CI 0.32-1.29).

Table 8.6: School attendance, HIV awareness, sexual behaviour and HIV infection among women aged 14-25

Variable	Proportion with high risk				aOR1	aOR2a		**aOR2b**		aOR3	P (Age Interaction)
	Not attending school		Attending school								
	N	%	N	%							
N	**380**		**623**								
HIV awareness											
Lack of HIV knowledge	188	49.5	373	59.9	1.13	1.17	0.83-1.66	**1.25**	**0.86-1.80**	1.22	0.967
Lack of household communication	174	45.8	347	55.7	1.22	1.41	0.99-2.01	**1.26**	**0.87-1.83**	1.24	0.556
Not had an HIV test	324	85.3	585	93.9	1.23	1.26	0.72-2.18	**0.78**	**0.44-1.41**	0.77	0.077
Holds stigmatising attitudes	221	58.2	399	64.0	0.93	0.94	0.66-1.34	**0.93**	**0.64-1.36**	0.89	0.729
Individual level sexual behaviour											
Ever had sex *	364	95.8	342	54.9	0.76	0.73	0.60-0.89	**0.95**	**0.77-1.17**	0.95	0.199
Unwanted sex at debut $	167	45.9	179	52.3	1.04	0.91	0.62-1.34	**0.99**	**0.66-1.48**	0.97	0.323
More than 2 partners during lifetime	197	51.8	110	17.7	0.59	0.54	0.37-0.79	**0.69**	**0.46-1.04**	0.69	0.233
More than 1 partner during previous year	25	6.6	32	5.1	0.89	0.82	0.40-1.69	**0.92**	**0.43-1.96**	0.92	0.166
Ever had a child	275	72.4	87	14.0	0.17	-	-	**0.16**	**0.11-0.24**	0.16	0.001
14-19years	236/320	73.8	72/576	13.0	0.12	-	-	0.12	**0.08-0.17**	**0.12**	-
20-25years	39/60	65.0	12/47	25.5	0.24	-	-	**0.22**	**0.07-0.66**	0.22	-
Continuing sexual activity (opp. 2° abstinence) $	327	89.8	309	90.4	1.22	1.14	0.60-2.15	**0.78**	**0.40-1.54**	0.79	0.052
Partnership characteristics *(N)*	**354**		**346**								
Intercourse > 5 times in past year	269	76.0	195	56.4	0.60	0.53	0.35-0.81	**0.55**	**0.36-0.85**	0.57	0.148
Male partner more than 3 yrs older than female	208	63.4	147	44.7	0.53	0.46	0.31-0.69	**0.51**	**0.34-0.78**	0.52	0.273
Only male partner provides resources	244	68.9	256	74.0	1.07	1.29	0.83-2.00	**1.27**	**0.80-1.99**	1.26	0.681
Never used a condom in the last year	186	52.5	153	44.2	0.63	0.64	0.43-0.94	**0.62**	**0.41-0.93**	0.60	0.624
No condom use at last sex	271	76.6	223	64.5	0.65	0.64	0.42-0.98	**0.71**	**0.45-1.11**	0.69	0.327
HIV *(N)*	**325**		**555**								
HIV Positive	58	17.9	50	9.0	0.84	0.79	0.46-1.36	**0.77**	**0.43-1.38**	0.77	0.470

aOR1 adjusted for age, aOR2a adjusted for age, confounders and household wealth, aOR2b adjusted for age, confounders, household wealth and childbirth, aOR3 adjusted for age, confounders, wealth, childbirth and temporary migrancy. Missing data was on age of sexual debut (7) and age difference between partners (45). * Effect estimate was Hazard ratio from Cox regression. ^ Adjustment for village pair only. $ Only among those who had previously had sex. Underlined when likelihood ratio test $p<0.05$, or when $0.05<p<0.10$.

Without adjustment for ever having had a child, but adjusted for other confounders, those in school were significantly less likely ever to have had sex. However, this effect was almost entirely removed after adjustment for ever having had a child. Individuals attending school were less likely to report multiple partners during their whole life with borderline significance, even after adjustment for previous childbirth. Individuals attending school were much less likely to have had a child. This effect was particularly strong among the younger group, but was also significant in the older group.

Sexual partnerships reported by those in school were significantly less likely to be characterized by a high frequency of sexual activity. Females attending school were slightly more likely to be in relationships characterised by male to female resource flow, although this relationship was not significant. However, those in school were significantly less likely to report partnerships characterized by the male being more than three years older than the female, and to report never having used a condom during the past year. There was also some evidence that lack of condom use at last sex was less common among those in school.

However, there was little evidence of an association between school attendance and prevalent HIV infection.

There was also little evidence to suggest that the associations between school attendance and HIV risk characteristics were confounded by temporary migrancy, or that any of the factors on the causal pathway mediated the associations noted with partnership level characteristics.

2) Temporary migrancy

Of relationships examined between temporary migrancy and risk characteristics, 10/16 odds ratios suggested lower risk among those who were temporary migrants, including all of the HIV awareness variables. The 95% confidence intervals around six of the odds ratios suggested statistically significantly lower risk characteristics among temporary migrants, with one significant association suggesting higher risk among temporary migrants. Additional adjustment for childbirth had little effect on the magnitude of measures of effect for most variables.

Temporary migrants were significantly more likely to have knowledge about HIV, to communicate with parents or other household members about sex/sexuality and to hold tolerant attitudes towards people living with HIV/AIDS. They were also more likely to have had an HIV test with borderline significance.

Temporary migrants were less likely to report sexual debut at which they had not wanted to have sex. There was little evidence of an association between temporary migrancy and lifetime number of partners among the whole group, but there was a statistically significant age interaction with this factor. Among the young group, non-migrants were as likely as migrants to have had multiple partners, while among the older group migrants reported more partners than non-migrants. There was also a statistically significant age interaction with ever having had a child. Among the younger age group there was little evidence of a relationship between migrancy and childbirth, while among the older group migrants were less likely to have had a child.

Table 8.7: Temporary migration, HIV awareness, sexual behaviour and HIV infection among women aged 14-25

Variable	Proportion with high risk Temporary Migrant				aOR1	aOR2a		aOR2b		aOR3	P (Age Interaction)
	No		Yes								
	N	%	N	%							
N	**896**		**107**								
HIV awareness											
Lack of HIV knowledge	519	57.9	42	39.3	0.53	0.51	0.33-0.78	0.51	0.33-0.78	0.51	0.694
Lack of household communication	477	53.2	44	41.1	0.68	0.68	0.44-1.04	0.67	0.43-1.02	0.67	0.340
Not had an HIV test	821	91.6	88	82.2	0.57	0.58	0.32-1.03	0.54	0.30-0.98	0.54	0.492
Holds stigmatising attitudes	583	65.1	37	34.6	0.31	0.31	0.20-0.48	0.31	0.20-0.48	0.31	0.896
Individual level sexual behaviour											
Ever had sex *	616	68.8	90	84.1	1.19	1.21	0.97-1.53	1.23	0.98-1.55	1.23	0.408
Unwanted sex at debut $	312	50.7	34	37.8	0.64	0.61	0.37-0.98	0.61	0.38-1.00	0.61	0.223
More than 2 partners during lifetime	259	28.9	48	44.9	1.25	1.24	0.78-1.98	1.29	0.80-2.06	1.27	0.025
14-19 years	92/579	15.9	6/42	14.3	0.72	0.62	0.23-1.64	0.52	0.19-1.44	0.48	-
20-25 years	167/317	52.7	42/65	64.6	1.57	1.64	0.92-2.93	1.86	1.03-3.37	1.90	-
More than 1 partner during previous year	50	5.6	7	6.5	1.10	1.11	0.48-2.57	1.13	0.48-2.62	1.13	0.736
Had a child	311	34.7	51	47.8	0.86	-	-	0.81	0.49-1.33	0.73	0.009
14-19 years	85/579	14.7	12/42	28.6	2.16	-	-	1.57	0.68-3.60	0.78	-
20-25 years	226/317	71.3	39/65	60.0	0.55	-	-	0.53	0.30-0.96	0.57	-
Continuing sexual activity (opp. 2° abstinence) $	552	89.6	84	93.3	1.58	1.42	0.57-3.51	1.31	0.52-3.26	1.28	0.266
Partnership characteristics *(N)*	**609**		**91**								
Intercourse > 5 times in past year	389	63.9	75	82.4	2.17	2.23	1.24-4.02	2.24	1.24-4.03	2.11	0.789
Male partner more than 3 yrs older than female	297	52.2	58	65.9	1.57	1.56	0.95-2.53	1.60	0.98-2.61	1.54	0.172
Only male partner provides resources	437	71.8	63	69.2	0.96	0.87	0.52-1.45	0.86	0.52-1.44	0.87	0.430
Never used a condom in the last year	303	49.8	36	39.6	0.64	0.61	0.38-0.97	0.61	0.38-0.97	0.58	0.134
No condom use at last sex	437	71.8	57	62,6	0.57	0.54	0.33-0.87	0.54	0.33-0.88	0.53	0.335
HIV *(N)*	**798**		**82**								
HIV Positive	96	12.0	12	14.6	0.96	0.99	0.50-1.95	0.99	0.50-1.95	0.99	0.209

aOR1 adjusted for age, aOR2a adjusted for age, confounders and household wealth, aOR2b adjusted for age, confounders, household wealth and childbirth, aOR3 adjusted for age, confounders, wealth, childbirth and attendance at school. Missing data was on age of sexual debut (7) and age difference between partners (45). * Effect estimate was Hazard ratio from Cox regression. ^ Adjustment for village pair only. $ Only among those who had previously had sex. Underlined when likelihood ratio test $p<0.05$, or when $0.05<p<0.10$.

Female temporary migrants were significantly more likely to report partnerships characterised by a high frequency of sexual intercourse. Migrants were also more likely to report partnerships in which there was a greater than three year age-gap to an older male. However, sexually active temporary migrants were significantly less likely to report never having used a condom and lack of condom use at last sex.

There was no evidence of an association between current temporary migrancy and prevalent HIV infection among females.

There was some evidence to suggest that differences in HIV awareness profiles between migrants and non-migrants explained part of the difference in the partnership sexual behaviour characteristics seen between the groups. Including terms for all HIV awareness variables in the model had a moderate impact on the odds ratio associated with sexual intercourse frequency (aOR 2.48 95% CI 1.36-4.53) and both of the condom use variables (never having used a condom in the past year, aOR 0.71 95% CI 0.41-1.45; no condom use at last sex aOR 0.62 95% CI 0.37-1.03).

Finally among the females, there was some evidence of interaction between schooling and migrancy in their relationship with some outcomes (Table 8.5). The significant association between migrancy and unwanted sex at debut in the whole group was not seen among the non-migrants (aOR 1.19 95% CI 0.76-1.87), but was seen among the smaller group of migrants (aOR 0.31 95% CI 0.09-1.12). There appeared to also be a significant interaction between schooling and migrancy with respect to secondary abstinence.

However, there were very low numbers in some groups when considering the migrant strata and as such the test was unstable.

8.03.3 Interaction and mediation between household wealth and individual socioeconomic roles in their association with HIV risk outcomes

(a) Males

There was little evidence from this study to suggest interaction between schooling or migrancy and household wealth in their associations with HIV risk characteristics (Table 8.8). Only one interaction test was significant at the $p<0.05$ level, in the relationship between schooling and wealth, for the outcome variable “lack of HIV knowledge”. Among the very poor group, those in school were less likely to report a lack of HIV knowledge (aOR 0.47 95% CI 0.23-0.98), while among the remaining group there was no association between school attendance and HIV knowledge (aOR 0.98, 95% CI 0.62-1.55).

Table 8.8: Likelihood ratio tests for interaction between household wealth and migration / school attendance in their association with HIV risk outcomes

Variable	Males		Females	
	Schooling	Migration	Schooling	Migration
	p	p	p	p
HIV awareness				
Lack of HIV knowledge	0.005	0.481	0.826	0.596
Lack of household communication	0.736	0.300	0.148	0.403
Not had an HIV test	0.519	0.397	0.971	0.027
Holds stigmatising attitudes	0.242	0.056	0.167	0.710
Individual level sexual behaviour				
Ever had sex *	0.354	0.463	0.913	0.132
Unwanted sex at debut $	0.534	-	0.392	0.091
More than 3 partners during lifetime	0.474	0.366	0.464	0.088
More than 1 partner during previous year	0.995	0.818	0.838	0.889
Childbirth	-	-	0.658	0.413
Continuing sexual activity (opp. 2° abstinence) $	0.401	0.550	0.657	-
Co-factor				
Circumcision	0.698	0.081	-	-
Partnership characteristics				
Intercourse frequency in past year	0.535	0.805	0.736	0.973
Male partner more than 3 yrs older than female	0.914	0.376	0.769	0.615
Only male partner provides resources	0.362	0.110	0.964	0.221
Never used a condom in the last year	0.614	0.224	0.397	0.069
No condom use at last sex	0.305	0.765	0.693	0.316
HIV				
HIV Positive	0.355	0.517	0.908	0.539

Likelihood ratio tests comparing models adjusted for all confounding factors. * Cox regression. $ Only among those who had previously had sex. Underlined when likelihood ratio test $p<0.05$, or when $0.05<p<0.10$.

Table 8.9: Mediation of the impact of household wealth on HIV awareness, sexual behaviour and HIV infection by individual socioeconomic roles among individual of both sexes aged 14-25 years

Variable	Males				Females			
	aOR comparing "poor, but a bit better off" & "doing OK" with very poor"	+ adj Schooling	+ adj Migration	+ adj schooling and migration	aOR comparing "poor, but a bit better off" & "doing OK" with very poor"	+ adj Schooling	+ adj Migration	+ adj schooling and migration
HIV awareness								
Lack of HIV knowledge	0.97	0.97	0.96	0.96	0.90	0.89	0.89	0.89
Lack of household communication	0.75	0.75	0.73	0.74	0.76	0.75	0.76	0.75
Not had an HIV test	1.48^	1.47^	1.47^	1.46^	0.68	0.69	0.68	0.69
Holds stigmatising attitudes	0.94	0.95	0.91	0.92	0.95	0.95	0.94	0.95
Individual level sexual behaviour								
Ever had sex *	0.93	0.93	0.92	0.93	1.08	1.08	1.08	1.09
Unwanted sex at debut $	1.03	1.03	1.03	1.03	1.21	1.21	1.23	1.23
More than 3 partners during lifetime	1.08	1.09	1.09	1.10	1.25	1.28	1.25	1.27
More than 1 partner during previous year	1.08	1.09	1.08	1.09	1.53	1.54	1.53	1.54
Had a child	-	-	-	-	0.75	0.83	0.75	0.83
Continuing sexual activity (opp. 2° abstinence) $	1.24	1.21	1.26	1.23	0.86	0.87	0.85	0.86
Co-factor								
Circumcision	1.30	1.28	1.30	1.28	-	-	-	-
Partnership characteristics								
Intercourse frequency in past year	1.06	1.06	1.04	1.05	0.70	0.67	0.70	0.68
Male partner more than 3 yrs older than female	0.82	0.82	0.82	0.82	1.01	1.05	1.01	1.05
Only male partner provides resources	1.01	1.01	1.02	1.02	0.84	0.82	0.84	0.82
Never used a condom in the last year	1.40	1.40	1.39	1.39	0.78	0.80	0.79	0.81
No condom use at last sex	1.09	1.09	1.09	1.08	0.76	0.77	0.76	0.78
HIV								
HIV Positive	1.28^	1.45^	1.30^	1.46^	1.08	1.10	1.08	1.10

aOR adjusted for age, confounders, and, among females only, for childbirth. Missing data was on age of sexual debut (15 males, 7 females) and age difference between partners (42 males 45 females). * Effect estimate was Hazard ratio from Cox regression. ^ Adjustment for village pair only. $ Only among those who had previously had sex

Having established that there was not interaction between household wealth and either socioeconomic role for outcome variables among males, the final stage of the analysis sought to identify whether these roles mediated any of the association between wealth and HIV risk outcome. Table 8.9 shows how the odds ratio for each HIV risk characteristic associated with the "poor, but a bit better off" and "doing OK" groups together compared to the "very poor" group changed when variables coding for attendance at school, temporary migrancy or both role variables were added to the models. Generally there was very little change in the odds ratio associated with household wealth for each risk characteristic upon inclusion of the schooling and temporary migrancy variables to the model. In only one case did the odds ratio change by more than 0.1. The odds ratio associated with prevalent HIV infection rose from 1.28 to 1.45 after inclusion of the school attendance variable (95% CI 0.67-3.13). This suggested that were it not for the fact that young men from wealthier households were slightly more likely to be attending school, they would have had an increased risk of HIV infection relative to young men from poorer households, although this relationship was still not statistically significant.

(b) Females

The pattern among females was similar to that for the males in that there was little evidence to suggest interaction between schooling or migrancy and household wealth in their associations with HIV risk characteristics (Table 8.8). Again, only one interaction test was significant at the $p<0.05$ level, this time in the relationship between temporary migration and wealth, for the outcome variable "never had an HIV test". On further inspection there were very low numbers in some cells when the relationship was stratified on wealth (only 2 temporary migrants had ever had an HIV test), and the interaction test was therefore unstable.

Finally, there was also very little evidence of a shift in the odds ratio for household wealth with relation to all HIV risk characteristics upon inclusion of terms for school attendance and / or temporary migrancy in the model, giving little evidence from this study to suggest that these roles mediated any associations between wealth and HIV risk.

8.04 Summary of key findings

Few outcome variables were associated with household wealth among either sex and where effects were seen these were generally modest. The exception to this was that among females, condom use was significantly more often reported by those from wealthier households. There was no association between household wealth and risk of HIV infection among either sex.

Among males, there was some evidence to suggest a lower HIV risk profile among those attending school or college. Attendance at school was associated with a lower number of sexual partners. Among those sexually active in the previous year, males attending school/college were less likely to report relationships characterised by male to female resource flow. Prevalent HIV infection rates were significantly lower among those attending school, particularly among non-migrants. Male temporary migrants communicated more with household members about sex/sexuality and held less stigmatising views. However, migrants reported sexual partnerships with a higher frequency of sexual intercourse and more often provided resources to the female partner than non-migrants. Among the younger but not the older males, there was some suggestion that HIV prevalence was higher among the small number of migrants.

Among females, there was some evidence to suggest lower HIV risk characteristics among those who were in school. Individuals attending school were less likely to report more than two partners during the lifetime. They were also much less likely never to have had a child. They reported relationships more often characterised by low frequency of sexual intercourse, a smaller age gap and condom use. Female temporary migrants had higher HIV awareness levels and were more likely than those who stayed at home to use condoms, despite more often being in relationships characterised by frequent sexual intercourse and wide age gaps. Among the older but not the younger group, female migrants reported higher numbers of sexual partners.

There was little evidence in this analysis for the association between household wealth or socioeconomic role and outcome variables consistently being modified by age. There was additionally little evidence to suggest that any associations between socioeconomic exposures and any outcome variables were modified or mediated by differences in variables postulated to lie on the causal pathway being adopted by young people.

Chapter 9. Sensitivity analysis of key findings

9.01 Introduction

The interpretation of epidemiological studies requires consideration of the potential role of bias in distorting the results. This study set out to assess whether socioeconomic factors were associated with HIV awareness, sexual behaviour and risk of HIV infection. In interpreting the results, assumptions must be made about the robustness of the statistical techniques used, the validity of the collected data and the representativeness of the study sample. This chapter explores the sensitivity of the findings presented in this thesis to some of these assumptions.

9.02 Objectives of chapter

1) To calculate measures of the association between socioeconomic factors and three key HIV risk characteristics (lifetime number of partners, condom use at last sex, HIV prevalence),

    a) Using statistical techniques that account for clustering of outcomes by household and individual

    b) On a dataset in which missing outcome data on individuals who were eligible but not included in the final sample were replaced with simulated data on these individuals representing extreme biases in their outcome profile (high risk / low risk)

c) From a dataset where data on outcomes were simulated to reflect potential impacts of social desirability bias on reporting of sexual behaviour

d) Using a different method to classify households of different wealth

e) Excluding data on HIV status that were felt potentially to be unreliable.

2) To compare the results of these analyses with the original results presented in this thesis, to assess the sensitivity of the reported findings to assumptions about the validity of the study methods.

9.03 Methods

9.03.1 Clustering

It was possible that the analyses presented in this thesis may have been prone to error due to clustering of outcome variables. Throughout the analysis measures of effect and their 95% confidence intervals were computed using standard methodology that assumes sampling independence between analytic units. However, in this study households were the unit of sampling, yet analyses considered individuals, and, in some cases, sexual partnerships as the primary analytic units. Individuals from the same household, and sexual partnerships reported by the same individual, may be more alike than would be expected if these analytic units had been sampled directly. This phenomenon is known as *clustering.* There is the potential for two main impacts of this clustering. Firstly, measures of effect may be inappropriately estimated since standard methods give equal weight to multiple individuals from the same household (or partnerships reported by the same individual), yet if clustering does exist such records should be given relatively less weight. Secondly, ignoring clustering leads to an overestimation of the precision of presented measures of effect, and thus to the possibility for type I error. Type I error occurs when it is concluded that there is a difference in outcomes between exposure groups when in reality there is not. Analyses of key variables presented in this thesis computed point estimates of effect in the form of odds ratios with 95% confidence intervals. If standard errors were significantly underestimated, these confidence intervals may have been too narrow.

A number of methods are available that take account of correlation within sampled clusters in the computation of measures of effect and their standard errors. Two such approaches were used to recompute the major measures of effect and 95% confidence intervals presented. The first approach was the use of general estimating equations with computation of robust standard errors. This approach takes account of correlations within clusters in computing the odds ratio and computes standard errors by using the observed variability in the data rather than variability predicted by an underlying distribution. The population averaged *"xtlogit"* command in Stata v.8 was used for this purpose, specifying an *exchangeable* correlation matrix. Secondly, a random effects approach was used. Random effects models include the variation between clusters explicitly in the likelihood calculation. Such models can thus additionally be used to model between cluster variance and to estimate the level of correlation within clusters (households or individuals in this case). The *"xtlogit"* command specifying random effects in Stata v.8 was used to operationalise this approach. In addition, however, the reliability of random effects models for logistic regression need to be checked since likelihoods are estimated from numerical approximations of extremely complex underlying probability distributions. This was done by comparing the relative difference of the estimated parameters when different numbers of cut-points were used to perform these approximations. The Stata command *"quadchk"* was used for this purpose. If parameter estimates are sensitive to differences in the number of cut-off points the results of random effects should be treated with caution.

9.03.2 Selection Bias

A number of eligible individuals were not successfully interviewed, not included in the analysis due to missing data, or did not provide a sample for HIV testing. If the outcome and exposure profile of individuals who did not feature in the analysis was the same as for those who were successfully included then the main impact of this on the study would have been a reduction in statistical power. However, if individuals who were interviewed had a different outcome profile compared to individuals who remained uninterviewed, then selection bias might have distorted the results of the study.

Individuals were not included in the final sample for a number of reasons, although the most important was that eligible individuals were not found despite repeated visits to the home and thus did not have a young person's questionnaire available (see Chapter 5). Exposure data, age (approximated by year of interview minus year of birth) and two external confounder variables (village of residence, number of co-resident 14-25 year olds) were recorded on the household questionnaire and thus data was available for most of the individuals not included. The composite temporary migrancy variable used for most of this analysis was not available for most uninterviewed individuals. However, data on whether each individual was currently sleeping at home was available and individuals could be assigned migrancy status on the basis of this variable alone meaning that some individuals who would have been classified as temporary migrants under the expanded definition would have been incorrectly classified as having remained at home.

Outcome variables were simulated for this group. In order to estimate the potential impact of selection bias, two extreme situations were considered for each exposure – outcome combination. In the first scenario, individuals who were not included in the final sample, or, in the case of HIV serostatus, did not have HIV serostatus data, were simulated as having extremely high HIV risk characteristics. In practice this meant setting all of these individuals to the high risk outcome (high number of partners, lack of condom use at last sexual intercourse, HIV positive). Since the majority of excluded individuals were in the older age group, as a rough approximation it was assumed that all these individuals were sexually active, and condom use was simulated for one partner for each of these individuals. Following this, the odds ratios for each exposure-outcome combination were recomputed using the full model without inclusion of the two external confounders for males (orphanhood status and place of birth) and three for females (childbirth) on which no data was available. The exclusion of these variables had limited impact on the odds ratios from the main models.

In the second scenario, all individuals were assigned to the low risk status (low number of partners, condom use at last sex, HIV negative), and logistic regression models were run again.

9.03.3 Social desirability bias

Individuals may have misreported their own sexual behaviour since some behaviours, such as having multiple sexual partners or not using condoms, may have been perceived by respondents as being "unacceptable" to the interviewer or more generally in society. This type of social desirability bias has been reported in a number of studies (214, 215). If misreport of sexual behaviour was randomly distributed across exposure groups then the measures of effect presented would most likely be underestimates of the true association between socioeconomic factors and HIV risk characteristics. However, if individuals from different groups differentially misreported their sexual behaviour characteristics then information bias might have distorted the results presented in more complex ways. Misreporting on the basis of social desirability may be particularly important source of bias in studies of behaviour among different social groups.

Descriptive data providing an insight into misreport of sexual behaviour were provided in two ways. Firstly, the prevalence of HIV infection among individuals of either sex reporting never having had sexual intercourse was calculated. Secondly, the total number of reported sexual partnerships during the past year was computed for both sexes (216). Sexual partnerships with individuals not within the age group under study were excluded. This number was divided by the response rate for each sex. Comparison of these numbers lends insight into differential misreporting of sexual behaviour by males and females. Differences may also be seen if individuals of either sex form significant numbers of sexual partnerships with individuals from outside the study area. Two further analyses were performed considering firstly only those individuals who reported currently sleeping in the home, and secondly, only those individuals who reported having stayed in the home for all 12 months of the previous year, since these individuals were the least likely to have formed relationships with individuals away from the study site.

For the sensitivity analysis it was hypothesised that some individuals may have been more prone to misreport sexual behaviours than others. Specifically, it was hypothesized that individuals from the wealthiest households, those remaining in school and those who were temporary migrants may be more likely to over-report safe sexual behaviours since these individuals had higher levels of HIV awareness and may have been particularly prone to social pressure in reporting. In each case two simulations were considered. In the first simulation, 5% of individuals in the above exposure groups who reported low-risk responses for the two sexual behaviour outcome variables were simulated to have misreported this fact. A random 5% of records were altered to reflect this for each of the two sexual behaviour outcome variables. In the second simulation, differential misreporting as described above was simulated at 20%. Odds ratios and 95% confidence intervals were recomputed for both sexual behaviour outcome variables after these simulations.

9.03.4 Different measure of household of wealth

Chapter 6 described the generation of three measures of household wealth from the available study data, though only one was used throughout the thesis. The study may have misrepresented the association between relative household wealth and risk of HIV infection if the method chosen was inaccurate. There were both methodological and conceptual differences between the different methods used to generate an index of household wealth, and both may have contributed to different ranking of households by these methods. Closer examination of these issues is held in the discussion.

The measure of association between the three outcome variables and household wealth was recomputed when the PCA method was used to generate the index of wealth, and households were divided into terciles on the basis of this measure. This measure was used since it had the lowest level of agreement with the PWR measure that was used throughout this thesis.

9.03.5 Validity of HIV testing

Laboratory errors described in chapter 7 of this thesis meant that a significant proportion of oral fluid (OMT) samples were processed beyond the 21-day cut off recommended by the manufacturers. These samples may have been less accurate in their determination of HIV status than those samples processed appropriately.

Measures of association between each of the socioeconomic exposure variables and HIV infection were calculated using only HIV serostatus data available from individuals whose samples were processed within 21 days of collection, as required in the manufacturers guidelines. Differences in the results seen might have had a number of origins. Firstly, sampling error might have led the point estimates from the two different analyses to be different. The results of analyses conducted only on samples processed within 21 days had less statistical power to detect associations and 95% confidence intervals were expected to be wider. Secondly, samples analysed outside 21 days may have been more likely to be inaccurate in their assessment of HIV status and this may have affected the relationship between HIV and outcomes. This would have been compounded if samples that were analysed outside 21 days were not a random sample of all samples, but some individual traits associated with the exposures under study were also associated with whether or not samples were analysed within the manufacturer's guideline timelines.

In addition, during data cleaning it was found that OMT samples from one fieldworker had a much greater HIV prevalence than those from all other fieldworkers. Although the reasons for this remained unknown, it was felt that these results needed to be treated with some caution. Measures of effect were re-calculated after the results from this fieldworker were excluded.

9.04 Results

The primary study findings are summarized in bold in tables 12.1 and 12.2, with odds ratios computed after each of the above simulations held below.

9.04.1 Clustering

The results of the main analyses re-computed using a general estimating equations approach with computation of robust standard errors to account for household (all outcomes) and individual (condom use) clustering are held in Tables 12.1 and 12.2. For both sexes it was found that there were only very small shifts in either the point estimates or the width of confidence intervals when this approach was used.

Table 9.1: Results of sensitivity analysis of key findings presented in this thesis for males

Indicator	N	Household Wealth (Comparing "Doing OK" with "Very poor" group)		Attendance at School		Temporary migrancy	
		Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval
Lifetime partners > 3	916	1.06	0.61-1.84	0.67	0.45-0.99	1.15	0.74-1.78
GEE estimate (Household clustering)	916	1.07	0.61-1.88	0.66	0.44-1.00	1.13	0.74-1.74
Selection Bias (Low)	1235	1.12	0.76-1.43	0.74	0.52-1.05	0.47	0.33-0.67
Selection Bias (High)	1235	1.02	0.66-1.56	0.75	0.54-1.06	2.60	1.86-3.65
Social desirability bias (5%)	916	1.17	0.69-1.99	0.76	0.52-1.12	1.26	0.82-1.94
Social desirability bias (20%)	916	2.08	1.24-3.51	1.19	0.81-1.74	1.87	1.21-2.88
PCA indicator of household wealth	909	1.36	0.89-2.08				
Lack of condom use at last sex within partnership	748	0.71	0.41-1.22	0.84	0.56-1.27	0.74	0.49-1.12
GEE estimate (Household clustering)	748	0.77	0.43-1.37	0.78	0.50-1.24	0.72	0.45-1.14
GEE estimate (Individual clustering)	748	0.74	0.42-1.32	0.82	0.52-1.28	0.70	0.44-1.10
Selection Bias (Low)	842	0.90	0.55-1.46	0.82	0.56-1.20	0.30	0.21-0.44
Selection Bias (High)	842	0.70	0.39-1.21	0.83	0.53-1.31	1.63	1.06-2.51
Social desirability bias (5%)	748	0.74	0.43-1.25	0.82	0.55-1.22	1.79	1.52-1.19
Social desirability bias (20%)	748	0.99	0.58-1.09	1.16	0.77-1.74	1.06	0.09-1.61
PCA indicator of household wealth	517	0.91	0.55-1.50				
HIV positive	763	1.44	0.50-4.12	0.27	0.11-0.66	1.17	0.49-2.76
GEE estimate (Household clustering)	763	1.43	0.50-4.07	0.28	0.09-0.86	1.16	0.41-3.27
Selection Bias (Low)	1082	1.23	0.44-3.38	0.29	0.12-0.72	0.59	0.26-1.32
Selection Bias (High)	1082	0.95	0.61-1.48	0.75	0.52-1.09	4.41	3.17-6.13
PCA indicator of household wealth	757	1.19	0.53-2.69				
HIV test validity – 21 days	462	0.99	0.26-3.72	0.34	0.11-1.01	1.69	0.58-4.89
HIV test validity – excluded fieldworker	709	0.99	0.23-4.28	0.21	0.06-0.70	0.70	0.22-2.28

Table 9.2: Results of sensitivity analysis of key findings presented in this thesis for females

Indicator	N	Household Wealth (Comparing "Doing OK" with "Very poor" group)		Attendance at School		Temporary Migrancy	
		Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval	Odds Ratio	95% Confidence Interval
Lifetime partners > 2	**1003**	**1.28**	**0.86-1.79**	**0.69**	**0.42-1.04**	**1.27**	**0.80-2.04**
GEE estimate (Household clustering)	1003	1.27	0.75-2.15	0.69	0.40-1.03	1.28	0.82-2.00
Selection Bias (Low)	1234	1.12	0.72-1.74	0.53	0.37-0.76	0.51	0.34-0.75
Selection Bias (High)	1234	0.94	0.65-1.42	0.74	0.54-1.03	3.00	2.10-4.24
Social desirability bias (5%)	1003	1.49	0.92-2.43	0.74	0.51-1.07	1.38	0.87-2.19
Social desirability bias (20%)	1003	2.61	1.63-4.18	1.21	0.84-1.73	2.15	1.35-3.42
PCA indicator of household wealth	994	0.94	0.62-1.43				
Lack of condom use at last sex within partnership	**700**	**0.48**	**0.28-0.82**	**0.69**	**0.44-1.07**	**0.53**	**0.32-0.87**
GEE estimate (Household clustering)	700	0.50	0.28-1.28	0.70	0.45-1.09	0.53	0.32-0.87
GEE estimate (Individual clustering)	700	0.49	0.28-0.84	0.69	0.44-1.07	0.54	0.33-0.89
Selection Bias (Low)	867	0.68	0.70-1.29	0.55	0.35-0.79	0.25	0.17-0.36
Selection Bias (High)	867	0.,55	0.33-0.91	0.78	0.51-1.19	1.49	0.95-2.34
Social desirability bias (5%)	700	0.52	0.31-0.88	0.65	0.42-0.99	0.56	0.34-0.91
Social desirability bias (20%)	700	0.67	0.39-1.15	0.87	0.56-1.34	0.74	0.45-1.22
PCA indicator of household wealth	632	0.41	0.25-0.66				
HIV positive	**880**	**0.93**	**0.72-1.86**	**0.82**	**0.46-1.45**	**1.00**	**0.51-1.96**
GEE estimate (Household clustering)	880	1.03	0.52-2.07	0.80	0.42-1.52	1.04	0.52-2.09
Selection Bias (High)	1111	0.87	0.45-1.69	0.74	0.43-1.27	0.49	0.26-0.93
Selection Bias (Low)	1111	0.79	0.52-1.20	1.12	0.78-1.60	4.45	3.13-6.31
PCA indicator of household wealth	873	0.92	0.54-1.58				
HIV test validity – 21 days	583	0.79	0.34-1.84	0.72	0.36-1.46	1.03	0.45-02.40
HIV test validity – excluded fieldworker	809	0.62	0.26-1.46	0.75	0.40-1.40	0.65	0.28-1.52

Random effects models were also used to re-compute the measures of effect. However, in many cases, there was evidence that the model estimates might be invalid on the basis of the sensitivity check using different cut-offs to estimate the underlying distribution. Because of this these estimates were not considered further.

From this simulation it was felt that the results of the main study were unlikely to misrepresent the true situation because of correlation between outcome variables within clusters.

9.04.2 Selection Bias

There were moderate shifts in the odds ratios for the association between household wealth or school/college attendance and HIV risk characteristics after simulation of both extreme low and high risk characteristics among those who were not included in the final sample. In the majority of cases, these shifts would have altered the magnitude of the reported measure of effect, but not the broader conclusions about the presence or absence of an association being absent or present.

In contrast, relationships between temporary migrancy and outcomes were highly sensitive to the extreme simulations because may of those not included were migrants. In each case, a high risk simulation suggested that temporary migrancy was significantly associated with increased risk characteristics for every outcome, while under a low risk simulation, the opposite effect was seen for every outcome except HIV infection.

From this simulation it can be concluded that if there was a strong independent effect on HIV risk characteristics associated with non- inclusion in the final study sample, then the results presented with respect to the association between temporary migrancy and these characteristics had a high chance of being biased. It is not currently known if such an effect would be present in the study population, although future research might lend insight into this. Currently, this result does suggest that these results should be treated with additional caution and field surveys should be encouraged to use methods that ensure a high proportion of these individuals are included.

9.04.3 Social desirability bias

A total of 9 HIV infections (20.9% of all infections) among males and 15 HIV infections among females (13.9% of all infections) were detected among individuals who reported that they had never had sexual intercourse. The most extreme estimates accepted by UNAIDS suggest that up to 10% of all HIV infections may be acquired iatrogenically in some settings, although most estimates put the figure somewhat lower than this (65). These findings suggest that among both males and females there was some misreporting of ever-sexual activity.

Males reported a total of 774 partnerships during the previous 12 months. Where data was available, approximately 4.6% of partnerships were reported to be with females who were not within the age range 14-25 years (mainly partnerships involving females younger than this). The response rate among males was 74%. Calculating the estimated total number of sexual partnerships between 14-25 year olds as reported by males as (784-(0.046*74)]/0.74, gave an estimated 998 partnerships. For females, 702 partnerships were reported. However, 39% of those for which information was available involved male partners aged over 25 years. The response rate among females was 81%. The total number of partnerships for females was calculated as [702-(702*0.39)]/0.81, giving a total of 529 partnerships. The ratio between these two figures was 1.89. This suggested that males reported a higher number of sexual partnerships than females. Males may have been more likely to have had partnerships from outside the study area since they were more likely to be migrants. Consequently, the analysis was repeated twice, first considering only individuals of either sex who were currently sleeping at home, and secondly considering only individuals of either sex who reported staying at home for the whole year (12 months). When only those sleeping at home were considered the figures were 777 for males and 472 for females, giving a ratio of 1.68, somewhat lower than the previous estimate. When only individuals who were staying at home for all 12 months of the previous year were considered, the figures were 525 for males and 359 for females, giving a ratio of 1.46, suggesting that male migrants having partners from outside the study area may have led to overestimation of misreport in the original analysis. It was also possible that a high number of sexual partnerships reported by males were with women who had a very high number of partners such as commercial sex workers, but who may have been unlikely to be sampled in the survey. It was not possible to test this possibility with the current data set. From the available information, this analysis suggested that there was some evidence either that males relatively over-reported their number of sexual partners or females relatively underreported theirs.

The results of the social desirability analysis for the two sexual behaviour characteristics are shown in tables 9.1 and 9.2 In each case the simulations increased the numbers of 'exposed' individuals with the high risk outcome, and so the odds ratios tended to rise in each case. However, only under a strong assumption of 20% differential misreport among those in the exposed groups did the new estimate of effect in each case differ substantially from the original estimation. The results of the study were robust to a relatively small degree of biased misreport of outcome data on the basis of socioeconomic status.

9.04.4 Different measure of wealth

In each case the result of the association between household wealth and HIV risk outcome was different when wealth was estimated on the basis of the PCA method. No clear pattern emerged in the direction of this effect, although the shifts in the odds ratios were moderate to large in some cases. The results of the study, particularly in terms of the magnitude of any associations, were sensitive to which method was used to estimate household wealth and this needs to be considered in future research.

9.04.5 Validity of HIV testing

HIV prevalence was very similar among the samples analysed within 21 days (males 26/462, 5.6%, females 71/583, 12.2%) and those analysed outside the 21 day guideline (males 17/301, 5.7%, females 37/297, 12.5%).

No systematic reason linking the processing of samples within 21 days to sociodemographic characteristics of the individuals from which they came was observed or could be elucidated in discussions with lab staff. The processing time was determined by lab staff who were unaware of the characteristics of individuals from whom samples came. Within the data, having had a sample processed within 21 days was slightly more common among females than males (males 462/763, 60.6%, females 583/880, 66.3%, p=0.017). Having had a sample processed within 21 days was not associated with age, household wealth, attendance at school, temporary migration or ever having had sex among either sex.

In the analyses repeated using only samples processed within 21 days there were moderate variations in the measures of effect. It was difficult to interpret the findings seen in these analyses, but additional care was warranted in the interpretation of results relating to HIV status.

HIV prevalence was much higher among samples collected by one fieldworker For this fieldworker's samples, among males 19/54 (35.2%) were HIV positive, while among females HIV prevalence was 29/71 (40.9%). Thus, when this fieldworker was excluded from the analysis, HIV prevalence in the whole sample was much lower (27/709, 3.4% among males, 79/809, 9.8% among females). This may have been because of errors or because the fieldworker worked in a very high prevalence area since fieldworkers were assigned to work in specific locations. Having had a sample collected by this fieldworker was not associated with respondent's age, household wealth or socioeconomic role. Among males, but not females, those who had had a sample collected by this fieldworker were slightly more likely to report ever having had sex. When the main analyses were repeated excluding this fieldworker's samples, the odds ratio was lower in each case, and the drop was quite large in some cases. The conclusion that school attendance was associated with prevalent HIV infection was not affected by dropping this fieldworker's results from the analysis.

9.05 Summary of key findings

The results presented in this thesis were robust to intra-household and intra-individual level clustering in responses by participants. Results presented on the association between temporary migrancy and outcomes were highly sensitive to eligible individuals who were not included in the final study sample having a markedly different profile in terms of outcomes. The results for sexual behaviour had the potential to be influenced by biased misreport on the basis of socioeconomic factors if this had occurred at a relatively high level, and there was some evidence for misreport of sexual behaviour. The results were sensitive to the method used to generate an index of household wealth. Finally, additional caution was warranted in interpretation of the associations with HIV seroprevalence due to some mistakes in lab processing, the effect of which on accuracy in specimen processing was unknown, and because results from one fieldworker were suspect.

Chapter 10. Discussion

In 2001, the United Nations General Assembly Special Session on AIDS, perhaps the most important meeting yet to have been conducted on HIV/AIDS, recognised that "poverty, underdevelopment and illiteracy are among the principal contributing factors to the spread of HIV/AIDS" (217). That there is an association between poverty and HIV/AIDS is widely accepted in both academic and policy circles.

However, understanding the association between poverty and the risk of acquiring HIV infection is complex. This thesis has aimed to contribute in this regard by exploring the social epidemiology of sexual behaviour and HIV infection among unmarried young people in a rural area of South Africa. This closing chapter will begin with a summary of the main findings of the study. Following this, the chapter will discuss strengths and limitations of the research, reflecting on how these should inform interpretation of the results. The findings will then be discussed in relation to the gaps in the research literature identified earlier. The thesis will conclude with a brief discussion of the implications of the study findings for the control of HIV in sub-Saharan Africa.

10.01 Statement of principal findings of the research

This study demonstrated a high level of poverty among households in eight villages of a former bantustan in Limpopo Province, South Africa. As well as locally generated income, households were heavily dependant on migrant labour, state grants and small-scale enterprise to generate income, since local employment opportunities were few. The study also highlighted wealth inequality within the site. "Very poor" households were struggling to meet basic human needs such as food security and shelter, while only the wealthiest households, those that were "doing OK", had access to secure employment, more lucrative local business opportunities and higher standards of living.

Among a sample of 916 males and 1003 females, almost all young people of both sexes were attending school at age 14 years. This dropped steadily to less than 20% by age 25, including a high level of secondary school attendance at ages beyond 20 years, as well as some attendance at further education institutions. School attendance fell with age more quickly among females than males, with the difference probably driven to a great extent by pregnancy and childbearing. Females from "very poor" households were the least likely to be attending school. Relatively few individuals aged 20-25 years had completed the secondary school qualification. Among those who were not attending school, unemployment levels were high, with most young people looking for work and doing occasional jobs. Young people, including both students and those out of school, were increasingly likely with increasing age to be temporary migrants away from the rural home. At younger ages, migration was more common among individuals who had left school, but among those aged 20-25 years migrants were as likely to be attending school or college as not.

Lack of knowledge about HIV and holding stigmatising attitudes towards people living with HIV/AIDS (PWAs) were both widespread. Intra-household communication on sex and sexuality and having had an HIV test were rarely reported. Young people were more likely to have been sexually active or to have had multiple partners with increasing age. Males reported higher numbers of partners than females, though there was little difference between men and women in age at sexual debut or reported condom use. Condom use at last sex was reported for about one third of sexual partnerships. Sexual partnerships reported by young women usually involved males who were older than them and who they perceived to be providing them with resources, while young men reported partnerships with women who were of a similar age or younger. HIV prevalence was 5.6% among males and 12.2% among females, increasing over the age range. Individuals of either sex who held stigmatising attitudes towards PWAs were less likely to be HIV infected. Among females, increased risk of infection was associated with an increase in total number of reported partners.

In relation to the primary study hypotheses, household wealth was generally not associated with HIV risk characteristics among either males or females. An exception to this among males was that holding stigmatising views about people living with HIV was less commonly reported by young men who were from wealthier households. Among the females, those from wealthier households reported higher levels of condom use.

School attendance was associated with some lower HIV risk characteristics among both sexes. Among males, attendance at school or college was associated with a lower number of sexual partners, but not with condom use. There was some suggestion that relationships reported by young men attending school had a lower frequency of sexual intercourse and less often involved resource exchange. Males attending school or college were significantly less likely to be HIV positive, with some evidence that this effect was restricted to non-migrants. Among females, school attendance was also associated with a lower number of reported partners, even after adjustment for childbirth. Females attending school were less likely to have had sex or have had a child than those not attending school, although it seemed plausible that this association was partly due to reverse causality (e.g. individuals stopped attending school *because* they had had a child). Independent of childbearing, females in school reported relationships characterised by a lower frequency of sexual intercourse, smaller age gaps between partners and greater levels of condom use. Among females there was no association between school attendance and prevalent HIV infection.

Finally, there was some evidence that temporary migrancy was associated with lower HIV risk characteristics. Male temporary migrants were more likely to communicate with household members about sex / sexuality or to hold tolerant attitudes towards PWAs than those who stayed at home. The relatively small number of female temporary migrants reported higher levels of HIV awareness and condom use than those who were non-migrant, independent of household wealth or school attendance. However, there were also suggestions of other higher risk characteristics being more common among migrants. Among males, temporary migrants had relationships characterised by a higher frequency of sexual intercourse and greater levels of resource exchange. Among younger but not older males there was some suggestion of higher HIV prevalence among migrants, though this was based on small numbers. Among older, but not younger, females, migrants reported higher numbers of sexual partners.

10.02 Strengths and limitations of the research

Before discussing the wider implications of the results, a number of aspects of the design and conduct of the study need to be considered in interpreting the findings outlined above.

10.02.1 Temporality

A major limitation of the current study was the cross-sectional nature of the data. In cross-sectional studies it is not known whether measured exposure variables preceded the development of measured outcomes. Some relationships presented in this thesis warranted particular scrutiny in this regard.

The two socioeconomic role variables employed in this study measured the status of study subjects relating to the previous year. In neither case was the time point established at which these traits had been assumed. The measure of household wealth also reflected current status. Some outcome variables also measured recent behaviours, such as the characteristics of partnerships during the previous year, although the length of these partnerships was unknown. However, some outcome variables measured events that had occurred in the past, for example the acquisition of HIV infection; or characteristics that had been accumulated over some longer period, such as the lifetime number of sexual partners. This range of time dimensions in the data collected meant that it was particularly difficult to identify the temporal sequence for some relationships. For example, the association between socioeconomic roles and characteristics such as lifetime partners or HIV infection may be particularly difficult to interpret.

Generally, the impact of this problem in epidemiological cross-sectional studies is to misclassify time-dependent exposures and outcomes by the use of a classification captured at one point in time. If misclassification is not associated with exposure or outcome, then this would lead to an increased probability that true associations would not be captured by the study. This limitation in the study means that true associations between exposures and outcomes may not have been captured.

A potentially more serious problem is the situation where measured outcome variables at one point in time could have caused change in the exposure variable measured at the same point in time. This phenomenon, known as reverse causality, was a potential problem when examining the association between exposure variables and childbearing among young women. Previous studies have suggested that pregnancy and childbirth are leading causes of school leaving among females (21). In the data presented in this thesis, females between the ages of 17 and 22 were less likely than males to be attending school and it seems likely that at least part of this differential was the result of childbearing, although in South Africa teenage mothers may return to school once they have given birth (218), and young women may also leave school for reasons not related to childbirth. Childbearing among young women may also have reduced the chance of them migrating from the rural home or have affected the wealth of the household. The impact of fatherhood on male attendance at school in South Africa has not previously been reported, and no data was collected on that issue in this study.

The result of this reverse causality would be for an association to be seen between socioeconomic exposures and this outcome that is partly or wholly a result of this direction of causality, rather than that postulated in the conceptual framework (Chapter 2). As such, the association between attendance at school or temporary migrancy and childbearing should be interpreted with great caution. Previous childbirth is also likely to be associated with different levels of unprotected sexual activity, although the precise pattern is not simple to predict. Childbirth must be associated with some unprotected sexual activity. However, previous childbirth might also see reductions in sexual activity due to post partum abstinence, or an increased focus on contraception to prevent further pregnancy (219). Pregnancy and childbirth are also likely to affect HIV/AIDS knowledge, communication on sex / sexuality and HIV testing, through contact with health services. Because of these relationships, childbirth has the capacity to confound the relationships between socioeconomic exposures and some outcomes measured in this study. However, since childbirth was also highly correlated with these outcome variables, adjustment for childbirth might also cause an overadjustment of the true relationship between the socioeconomic exposure and sexual behaviour outcomes. In this thesis, relationships have been presented both with and without this adjustment. Greater weight in the final interpretation has been given to the potentially overadjusted figure, since this represents a more conservative approach. In practice, including a term for childbirth in the models had a moderate effect on many of the associations seen. This was strongest when attendance at school was the exposure variable under investigation. In some cases, for example the association between attendance at school and sexual debut, this adjustment attenuated a statistically significant association, although for most the effect was smaller than this.

Adjustment for previous childbirth does not take into account the potentially reverse causal effect of current pregnancy. It was likely that some young females were pregnant at the time of interview and not in school or migrant for this reason, but no data was collected on this, leaving the potential for further residual confounding. In summary, reverse causality due to pregnancy and childbirth was an important issue to deal with in this study. Some residual confounding by current pregnancy may have remained for some associations, although it is unlikely that this would have had a major impact on the results.

Reverse causality might also have been caused by illness induced by HIV infection. No data on current physical health was collected in this survey, but it was possible that some individuals were HIV positive and symptomatic with AIDS-associated illnesses. This would have the potential to impact both their socioeconomic roles and the wealth of the household. However, a strength of the current study in this regard was the focus on young people, in whom infection was likely to have been recent and who were accordingly less likely to be sick as a result of HIV infection.

Finally, issues of temporality in this study might also have weakened the capacity of the study to investigate mediation and interaction pathways, since each of the measured variables in the pathway indicated the status at a single point in time rather than their temporal sequence. The study might therefore have incorrectly concluded that there was no mediation or interaction between household wealth and individual socioeconomic role in their association with HIV risk outcomes. Future studies might use longitudinal data to overcome some of these issues although such studies often need to be very large and are thus expensive to do.

10.02.2 Study Power

The statistical power of the study to detect effects of moderate size was relatively low for many of the associations presented (see Table 5.4). Estimates of the magnitude of the associations presented in this thesis suggested that protective odds ratios lower than 0.5 were rare. In many cases the study did not have power to detect smaller associations than this. This was particularly a problem when prevalent HIV infection was the outcome under study, due to the low prevalence of this outcome. There were a number of reasons why effect sizes might have been quite small in this study. For example, differences in *absolute* wealth between those at the top and bottom of the relative scale used may have been quite small. This issue is discussed more later. Caution must be taken in interpreting negative as well as positive findings in the results of the analyses presented in this thesis. Future research projects will be strengthened by maximising sample sizes beyond that achieved in this study.

10.02.3 Potential mismeasurement of exposure variables

A strength of the research presented in this thesis was in the careful choice of a measure of relative household wealth. In addition, two socioeconomic roles relevant to the lives of young people were considered and conceptually linked to this. Previous empirical research in this area has generally not been guided by such conceptual frameworks. Particular issues relevant to considering socioeconomic exposures among young people were considered in this thesis. Thus, household wealth was judged to be a better marker of welfare than personal income, and attending school or college was considered an important socioeconomic role. Nevertheless, there may have been problems in the measurement or conceptualisation of the socioeconomic variables considered.

For this study, three methods for ranking household relative wealth were considered. The methods differed in two dimensions – that of the perspective they captured (the household being ranked or their neighbours) and the value ascribed to local (*emic*) or external (*etic*) perspectives on what constituted relative wealth. Each measure was designed to rank households from the poorest to the wealthiest, but these methodological differences meant that it was not surprising that the methods differed in their results. What was perhaps surprising was the magnitude of the differences noted, which meant that, as shown in the sensitivity analysis, the conclusions of the study might have differed slightly dependent on the method used to rank households. There follows a brief discussion on the similarities and differences, and strengths and weaknesses, of the approaches used.

There were many similarities in the techniques and the results obtained. Each was significantly associated with markers of asset ownership (including cars), housing quality, employment, human capital, food security and schooling of children. None was associated with the percentage of adults of working age in the household. Land tenure was less strongly associated with each of the measures. Household electrification was not significantly associated with the survey measures of wealth but was weakly associated with the PWR measure, while water availability was not associated with the PWR measure but was relatively weakly associated with both of the survey measures. It was perhaps unsurprising that some factors were less strongly associated with the indicators of relative wealth. The study area had limited agricultural potential and the economy is driven by cash income. Electricity was widely distributed within most households in six villages, but was rarely available in the other two. Similarly, despite a few households being wealthy enough to have access to yard water pipes, the majority were afflicted by poor water provision affecting the whole area. Thus, while water, electricity and land may be important for the generation of livelihoods it was perhaps not surprising that they were not strongly associated with relative household wealth assessed by any measure. Each of the techniques sought to identify households of relatively lower or higher wealth within a population who were all affected by these issues. Another common property of the three methods was that none explored intra-household resource allocation and so each implicitly assumed that the welfare level of all household members was equal. This assumption has been challenged in previous research, although most researchers continue to regard the household as a key unit of analysis (220).

Despite the similarities, the index of household wealth developed using PWR showed only modest agreement in its ranking of household wealth when compared to the survey measures. This was true even when themes from the PWR were used to inform the selection and weighting of the component indicators for the survey data. The PWR index of relative household wealth was chosen for use throughout the study for a number of reasons, including its high level of repeatability, local relevance and its independence from the measurement of the socioeconomic role variables. It is nevertheless useful to consider reasons why this technique gave different results to the survey methods. Two main potential reasons for the lack of agreement were identified; firstly, the techniques may have inaccurately assessed wealth due to inaccurate data collection or weighting; secondly, there may have been conceptual differences between the techniques so that in fact they measured different underlying concepts.

The survey methods attempted to maximise accurate reporting through collecting data on objective indicators, intensive fieldworker training and data quality monitoring, and stressing the importance of honesty to study participants. Nevertheless, reporting biases may be inherent to surveys in which individuals are asked to assess their own wealth. PWR partially accounts for such issues since information on household wealth is acquired from neighbours and is triangulated. However, it is also possible that PWR participants may provide invalid results for a number of reasons. Community residents may not know the wealth of other households in their village and can judge this only on the basis of what information is available. Households may conceal important information from their neighbours relevant to a true understanding of their wealth, and because of this PWR may best measure conspicuous consumption. PWR participants might also attempt to over or under estimate the wealth of certain households for reasons similar to survey participants. However, the high level of internal consistency for the household wealth scores obtained from three separate groups of PWR participants (ICC=0.81) was interesting in this regard. This finding differs from a previous report of low reliability for group-informant food-security ratings which are gathered in a similar manner to PWR (221). However, a number of the potential reasons for the low reliability reported by those authors have been addressed in this study, including the use of a relatively homogenous group of PWR participants, the presence of trained facilitators at all rankings, the use of locally generated definitions of poverty, and a focus on participation and consensus. The high level of internal consistency provided some evidence against major misreporting in the PWR data. However, PWR may also provide invalid results, but high levels of internal consistency, if participants, who were mostly poor women, ascribe a greater weight to certain dimensions of poverty (for example, being widowed) than would other groups in society.

The high level of migration also provided challenges to the estimation of household wealth for each of the techniques. Data on household members used in the PCA and combined PWR/survey method included information about migrants, including their employment and educational status and their ownership of assets. This was done since temporary migrants are important contributors to the rural economy in South Africa (10, 222). However, no information was available on levels of income remittance or connection between migrants and the household. PWR participants may not consider migrants so explicitly in their deliberations. They may not be considered part of the rural home. PWR participants may also be particularly uninformed about the wealth status of migrants from the rural home or their levels of income remittance. However, it is also possible that PWR participants had a more nuanced understanding of the role of migrants in generating household wealth than it was possible to capture from the available survey data.

Each of the indicators might not appropriately measure wealth even if the individual data were accurate. Principal components analysis is a complex statistical procedure that assigns weights to variables according to mathematical rules, while wealth ranking participants assess households in ways that are complex and non-transparent. The approach to PCA used in this thesis incorporated different facets of wealth and drew out the common underlying correlation between them. This may provide a more stable marker of wealth than individual markers used in isolation. This approach was designed to assess the overall welfare level of the household since it was felt this was most likely to influence the behaviour and roles of young people. However, the first principal component in the PCA explained only 22.7% of the total variance, suggesting that the factors included were not well correlated. While the technique was applied using a similar approach to previous authors (93), there also remain technical issues such as the inclusion of ordered categorical variables in PCA models that have not been well studied.

For the other measure of household wealth developed from survey data, information from PWR was used to inform the selection and weighting of indicators. This approach has intuitive appeal; however, there were limitations. These included that it was not always possible to directly map PWR statements to survey data and that the weighting system applied to the data was anyway arbitrary. Both of these limitations may have affected how well the combined index reflected underlying wealth. In addition, the indicators included information on both determinants and outcomes of wealth status, and mixed concepts such as standing wealth and current income. Despite the data from PWR suggesting that these were all perceived as important markers of wealth in the study community, it remains unknown if it is appropriate to combine information about these different properties to get an underlying measure of household welfare.

Finally, there was also room for differences in interpretation in the PWR method. Wealth ranking was conducted in *Sepedi*, using the question *"Mohloki wa mafelelo ke motho wa mohuta mang?"* to ask "what is a very poor person?" in order to start the ranking process. Many of the things identified by PWR participants as important in describing poverty in the setting resonated with the survey data. Nevertheless, the complexities of the ways in which PWR participants judged household wealth were inevitably lost in the generation of an index. One possibility is that PWR participants may have ranked households more directly on their current level of welfare than the survey based methods. In this respect, PWR may have been better correlated with markers of current income, although it was not possible to investigate this hypothesis in the study data.

Overall, careful consideration was given to the conceptualisation and measurement of household wealth used in this study. This work suggested that the marker adopted had good internal validity for measuring relative wealth as perceived by local people. However, the study also highlighted that this method gave results that differed from rankings of households by wealth achieved by two survey-based methods. There remain questions about the most appropriate measure of socioeconomic position to use in studies such as this that go beyond the remit of this thesis.

Socioeconomic role data were easier to capture, but might also have been mismeasured in this study. The socioeconomic role of attendance at school was considered. Those not attending school included those who were unemployed and not actively seeking work as well as members of the labour force. By grouping together those who were employed and unemployed, regardless of whether they were seeking or wanted to obtain work, the study assumed that this group were homogenous with regard their current "exposure". However, employed men and women tended to have slightly higher levels of educational attainment after adjustment for age (data not shown). This was judged appropriate, firstly because in this study setting unemployment was very high. Thus, those reporting current employment were nevertheless members of a high unemployment labour force. Additionally, much of the investigation in this thesis considered the impact of current school attendance on factors such as contact with HIV prevention campaigns and sexual network structure which would be common to those attending school.

Finally, a binary variable was used to capture whether individuals were temporary migrants or not. This approach meant that more detailed information on levels of mobility was not used in this thesis. This approach sought to identify migrants as a socioeconomic role group. Thus, individuals who may have made many trips, but for small lengths of time, were not considered to be migrants. Conversely, those who were away from the home for longer periods of time were considered as migrants. These individuals were felt to be more likely to have formed relationships with individuals from away from the study area or to have had the higher risk sexual behaviours assigned to migrants in previous studies (141).

In summary, this study carefully selected measures of socioeconomic status, but also highlighted complexities in this regard which need to be considered by researchers before conducting socioepidemiological research of this kind.

10.02.4 Potential mismeasurement of outcomes

Inaccurate measurement of outcome variables also had the potential to invalidate study findings. Previous work has highlighted complexities in the collection of data on sexual behaviour (215, 223, 224). A strength of this study was the consideration of such issues at the design phase. Fieldworkers were given extensive training, assessment and supervision, all interviews were guided by an extensive manual, most questions were objective rather than subjective, the interview was conducted in a safe place chosen by the respondent and the importance of honesty was stressed to participants (225). All interviews were conducted by females, since it was felt that female respondents may be particularly sensitive to interview gender, while male respondents may be less so. Interviewers were generally not known in the villages where they conducted interviews, but knew the study area well and sensitive questions were asked after rapport had been established between interviewer and respondent (215). Our approach was similar to that recommended for the collection of data on gender based violence in developing country settings (226). However, it was not possible to implement some recommendations for data collection on sexual behaviour, such as the use of non face-to-face interviewing techniques (225). There have been recent reports of a method of "informal confidential voting interview", in which respondents in a survey can conceal their responses to the most sensitive questions from the interviewer that may further reduce bias (215). This approach was not widely known when the current study was designed and was not adopted here.

One way of assessing the validity of reported data is to compare the descriptive findings on sexual behaviour patterns with similar studies. The data suggested many similar findings to other studies of sexual behaviour in South Africa. National studies disaggregated to provincial level have suggested a median age of first sexual intercourse between 16 and 18 years common to both sexes, a finding corroborated here (33, 42-46). This thesis also found that a significant minority of females begin sexual activity unwillingly, being forced to have sex the first time. This finding was also similar to other recent findings (59). The data also suggested that some individuals stopped sexual activity after previously being sexually active ("secondary abstinence"). However, levels of this were lower in this population than in previously reported national surveys of young people (36, 46).

As in previous surveys, our data suggested a higher rate of partner change among men than women (43, 45, 46, 49). Levels of partner change were difficult to directly compare with other studies due to heterogeneity in reporting, age distribution and marital status of study populations. Nevertheless, the data were similar to the LoveLife national survey of 15-24 year olds which reported 45% of males and 31% of females reporting more than three partners in their lifetime (46).

A high proportion of sexual partnerships were with individuals of the same age cohort. However, as age increased a higher proportion of partnerships comprised men who were more than three years the senior of their female partner, as in previous work (43, 46). Our definition of resource exchange was intentionally broad and was functionally different from "payment for sex" or "financial dependency". It was therefore quite different from previous reports in the South African literature (190). Finally, condom use at last sex within non-spousal partnerships was reported by a lower proportion of individuals than in recent national surveys (45, 46), though this was not unexpected given the rural study setting.

Overall, the patterns of reported sexual behaviour seemed highly plausible given previous literature in South Africa. The sensitivity analysis reported in Chapter 9 suggested some misreport of ever sexual activity, and differential misreport on numbers of sexual partners by men and women. The level of misreport suggested by these assessments was similar to or lower than that suggested by other recent studies (216). The sensitivity analysis also suggested that the results would be sensitive to differential misreport by individuals from different socioeconomic exposure classes, but it was not possible to examine whether this had occurred.

A further potential weakness of the data presented in this thesis was that some outcome variables were converted from naturally continuous form to binary data for use in logistic regression. This was done to facilitate the use of logistic regression for analysis and because cut-off values have been used in much previous work of this nature. Nevertheless, this loss of accuracy in outcome measurement meant that some associations in the data may have been missed if there was a small, but linear, association between variables, for example in the association between number of lifetime sexual partners and risk of HIV infection.

The addition of a biological marker to the behavioural assessment lends strength to the overall assessment (225), although the statistical power to evaluate associations was considerably less with this outcome than others. HIV prevalence was in a similar range to that suggested by other studies in the province (45, 46), and associated with number of lifetime partners among women, but not with recent condom use. Nevertheless, inaccuracy in determination of HIV status may have occurred. HIV infection in this study was assessed by ELISA assay of an oral fluid specimen. The OraSure HIV-1 Oral Specimen Collection Device has been licensed by the U.S. Food and Drug Administration (FDA) for collection of oral fluid specimens when used alongside the Oral Fluid Vironostika HIV-1 Microelisa System, since 1994 (http://www.fda.gov/bbs/topics/NEWS/ NEW00503.html) and is intended for use in subjects 13 years and older. Best available estimates of the sensitivity and specificity of the device / assay combination come from studies conducted in the U.S. and in Trinidad and Tobago. These studies report figures for both sensitivity and specificity in excess of 99%, from studies with 673 and 474 true positive results respectively (227, 228). However, there remains a dearth of good data on this device / assay combination from sub-Saharan Africa or its use under field conditions. Despite the lack of data, the technology has recently been used for a number of nationwide HIV prevalence surveys in South Africa (45, 46). A small sub-study conducted alongside this study identified a sensitivity of 93% (95% CI 66-100%) and specificity of 97% (95% CI 93-99%) for the assay compared to an established whole blood Elisa technique, although the sample size in this study was small (229). However, our data also provided some evidence to suggest a lower specificity and sensitivity among samples analysed over 21 days from collection, as is recommended by the manufacturer. In the sensitivity analysis it was shown that some of the results presented on the association between socioeconomic exposures and outcomes were changed by including only the samples processed according to the manufacturer's guidelines, although it was not possible to assess the reasons for this in detail.

In general, validated methods for the assessment of outcome variables were used in this study, as in similar studies identified in the literature review. Nevertheless, some inaccuracy in determination of outcome variables may have occurred, and this may have been most likely to increase the likelihood of the study finding no association between outcome and exposure variables.

10.02.5 Selection bias

As shown in Chapter 5, some potentially eligible individuals were not included in the final sample. If the association between exposures and outcomes among this group was similar to the whole sample, then the effect of this would only have been to reduce study power. However, if non-inclusion was associated with exposure status and independently with outcome status the results may have been biased.

There were a number of stages at which individuals might not have been included. Firstly, households might have remained unmapped and not had the chance to enter the study sampling frame. A small number of households did remain unmapped in two villages where this was checked. Nevertheless, the study mapping procedures were good and it is unlikely that this stage contributed significantly to selection bias.

Of sampled households, over 90% (1482/1640) were successfully enumerated. Of those that were not enumerated, nearly half turned out not to have any eligible individuals resident in them. A small number refused (15) and a small number of questionnaires were lost (10). In only 61 cases were dwellings thought to house some potentially eligible individuals but enumeration was not possible. It seems plausible that in at least some of these cases, inability to find a respondent in fact represented long-term absence of the whole household from the study area. In the remaining households, it also seems plausible that the average household size would have been quite small, since larger household size would have increased the chances of finding one household member to interview. On balance, it was felt that this source of non-inclusion was unlikely to have meant a large number of potentially eligible individuals were not included in the study sample.

Among enumerated individuals, a significant proportion were not interviewed. The data suggested that mobility was the major force driving non-inclusion in the study. The response rate among those currently sleeping in the home was much higher than among those not doing so. Further, inability to find a respondent was by far the most common reason for non-inclusion. Since migrants had a different socio-demographic profile to non-migrants, the study sample also under-represented those aged 20-25, males, those out of school, those with a secondary school certificate and those working for pay. It is possible that these individuals may also have independently had different characteristics in terms of HIV risk outcomes.

Our results in this regard confirmed the difficulty of conducting community based research, particularly in areas with high levels of migrancy. Studies that have examined the association between socioeconomic factors and HIV in other African settings have reported response rates between 56% and 90% (148, 150, 158, 165, 168, 173, 216, 230, 231). Some of these studies adopted procedures that included only individuals currently sleeping in the home, while a South African study designed to interview labour migrants at their rural homes achieved a response rate of 58% (29). The study employed a number of measures to maximise response rates, including high levels of field supervision, early morning and evening fieldwork periods and fieldwork over the Christmas period. Nevertheless, some individuals remained uninterviewed. Sensitivity analyses presented in Chapter 9 suggested that conclusions from the study with regard the association between temporary migrancy and outcomes might have been particularly susceptible to selection bias. However, the simulations presented in that chapter represented very extreme selection bias forces. Both approaches simulated situations in which a very strong force acted independently on all individuals who were not included in the final sample. This strong force was simulated to hugely increase the level of high risk or low risk outcomes among this group. This is an overly simplistic approach, and the distribution of outcomes in these simulations was unrealistic. Individuals were not included in the final analysis for a variety of reasons and these might have had different effects on the likelihood of selection bias occurring. More complex analyses might have attempted to simulate more realistic situations, including differences in the magnitude of bias, the relationship between exposures and outcomes among those not included, and covariance structures for this group. Nevertheless, it was useful to observe that associations between wealth or school attendance and HIV risk outcomes were relatively resistant to such bias, while the associations where temporary migrancy was the exposure were relatively sensitive to this.

For analyses where HIV was the outcome, sample sizes were again reduced since some individuals refused to provide a sample. The differences between those who provided an oral fluid sample and those who did not were similar to the differences between those interviewed and those not, although presumably quite different underlying dynamics were responsible for this. These individuals were more likely to be aged 20-25, male, out of school, holding a secondary school certificate and working for pay. Data in this case was also available on sexual behaviour characteristics. There was no difference in the proportion of individuals ever sexually active or having a non-spousal partner in the last year between those with and without an HIV test result. It was nevertheless possible that these individuals had higher (or lower) HIV risk characteristics and refused to provide a sample for reasons related to this.

In summary, selection bias had the potential to invalidate the conclusions of the study if those who were not included were different from those who were not in terms of their exposure and outcome profile. The results with regard temporary migrancy would have been particularly prone to this sort of bias.

10.02.6 Control of confounding

Confounding variables mask the true association between exposure and outcome. Regression techniques are used widely by epidemiologists to adjust analyses for confounding factors. Two issues relating to the control of confounding are relevant to a discussion of the interpretation of the findings presented here. Firstly, there was the possibility for residual confounding by variables that were not or were inaccurately measured. Secondly, where variables were controlled for in analysis, these procedures had the potential to overadjust the relationships of interest because proxy variables used to adjust for confounding concepts also represented variables that might have been on the causal pathway.

Age was strongly associated with many of the outcomes and with both socioeconomic role variables. As such, there was the potential for there to have been residual confounding by age. In the primary analyses, age was grouped into three year age bands. If smaller age bands had been used the power of models to detect associations would have been low, while larger age bands would have meant that there was the possibility for considerable variation within age bands in both exposures and outcomes leading to residual confounding. With the use of three year bands, residual confounding by age was felt unlikely to have the potential to seriously bias results.

Information on a number of additional potential confounders would have been useful. Cognitive ability may be associated with health seeking behaviours

(232). In this study it may also have influenced attendance at school or migration patterns, as well as sexual behaviour characteristics such as condom use. Direct data on this would have been difficult to collect. However, understanding of this issue might have been aided by data on reasons for leaving school or migration or on class marks at school.

As discussed earlier, data on current pregnancy would also have been a valuable addition to data on childbirth, since there remained the propensity for current pregnancy status to residually confound the relationship between attendance at school and sexual behaviours and risk of HIV infection, notwithstanding the issues of overadjustment described earlier.

The study would have benefited from data on the characteristics of sexual partners, including on their attendance at school, migration or wealth status and previous sexual behaviours or risk of HIV infection. Each of these might independently have influenced the reported patterns of sexual behaviour within partnerships, perhaps particularly among females, since male preferences in the timing and characteristics of sexual activity often prevail (51).

It would also have been useful to know more about the reasons socioeconomic roles were undertaken. School attendance might have been as much influenced by factors such as school availability, or the perception of the young person and their household head on the quality and usefulness of schooling, as by economic forces. Educational status and gender of the head of the household might also independently have influenced the attendance of young people at school (30), although these factors are also strongly associated with household wealth and are thus quite difficult to investigate as independent predictors of school attendance.

While some issues of lack of adjustment for confounding variables are outlined above, studies in which distal causal pathways are under investigation might also suffer from overadjustment. Distal factors might have a wide influence on downstream factors, which in turn might influence postulated outcomes through a variety of pathways. In analyses of socioeconomic exposures, it is usually of primary interest to assess the association between exposure and outcome through all relevant pathways, since alteration of the exposure level would likely affect all of these pathways simultaneously. Therefore, it is not, for example, usually appropriate to adjust the association between household wealth and HIV infection for differences in sexual behaviour between wealth groups unless non-sexual behaviour pathways are the explicit focus of attention. Such overadjustment has been a common feature of analyses on the social epidemiology of HIV (83, 153), although cross-sectional studies that have explicitly looked at the impact of overadjustment have generally not found large changes in reported odds ratios (147, 149).

In this study a small number of factors were adjusted for in all presented analyses, and it is worth considering whether this adjustment had the potential to overadjust analyses. The issue of adjustment for childbirth has been dealt with elsewhere, so the remainder of this section will focus on the other confounding variables adjusted for in analyses. Age was a clear confounding factor in that it did not have the potential to be influenced by either economic status or role, and was independently associated with most outcomes. Village was more complex in that current village of residence might have been influenced by wealth or income, since households of greater wealth may have been more likely to be resident in certain villages, perhaps those with better access to town or facilities. The process leading to orphanhood may also have affected household wealth and individual roles (in which case adjustment is appropriate). However, orphanhood might also have been more common among those of lower wealth due to the worse health profile of poorer households. It seemed less plausible that the number of co-resident young people was a pathway variable in any of the key relationships under study.

It was therefore theoretically possible that some analyses presented in this thesis were overadjusted. A pragmatic approach was taken throughout, in that all relationships have been presented both with and without adjustment for such factors. In most cases, the confounding influence of factors was relatively small.

10.02.7 Summary of validity issues

All observational epidemiological studies are constrained in their capacity to appropriately provide information on causal pathways. In this study, the cross-sectional nature of the data was problematic. The study was based within a longitudinal randomized trial of a socioeconomic intervention on sexual behaviour and HIV infection outcomes which will provide useful information in this regard. The study power was also low to detect associations with prevalent HIV infection. Efforts were made to deal with most other sources of potential bias in a manner comparable to or exceeding that adopted in most studies in this reason.

10.03 Implications of the study in relation to previous research

The literature review described in Chapter 3 of this thesis identified three major gaps in the literature on the social epidemiology of HIV infection in sub-Saharan Africa. This section will describe how the current study has contributed to understanding in these areas. Each will be tackled separately, although overlap clearly exists between the themes. Each of the sections below will include a brief discussion of contextual issues relevant to understanding the generalisability of the current study findings and will identify further research that could be conducted to refine hypotheses further.

10.03.1 The social epidemiology of sexual behaviour and HIV infection in South Africa

This study rejected the hypothesis of decreased HIV risk among those from households of greater wealth, although females of greater wealth were more likely to use condoms. The hypothesis of increased risk among temporary migrants of either sex was also largely rejected, although temporary migrants, particularly among age sub-groups, had some higher risk characteristics. The study provided some support for the hypothesis that young people attending school were at lower risk of HIV infection among males, and had lower behavioural risk profiles among both sexes.

While much literature has previously identified the potential role of poverty (233, 234) and labour migration (27, 141) in contributing to high risk behaviours there has previously been only limited empirical information available on the social epidemiology of HIV infection from South Africa. The literature review presented in this thesis identified only one national study, and some smaller studies focused on migrants in KwaZulu Natal, that presented data on the social distribution of HIV infection. Additionally, one national study presented the association between condom use and educational attainment and a recent study presented extensive data on the wealth distribution of sexual behaviours in KwaZulu Natal.

Methodological issues hampered interpretation of the national study, conducted in 2001. In that study, HIV was associated with borderline significance with a self-reported measure of income adequacy, such that those in the poorest groups had higher levels of infection, although this association was not seen when only the African population were considered (45). No objective marker of household wealth was available in that data set. In an unadjusted analysis of data from the same study, greater HIV prevalence was associated with higher educational attainment levels though this was difficult to interpret in the absence of adjustment for age, sex and setting. The report suggested that the effect of education was attenuated after adjustment for these factors, though few details were given. Studies focusing on labour migrants in KwaZulu Natal found no association between education and prevalent HIV infection. However, migrants and their partners had the highest levels of HIV infection in these studies (186, 196).

Regarding sexual behaviour, one national study found lower levels of condom use among female adults with less educational attainment in 1998 (56). A more recent study of sexual behaviour in KwaZulu Natal found limited evidence for an association between household wealth and a number of sexual behaviour characteristics among males, while among females a number of risky sexual behaviour traits, including lack of condom use, were associated with lower wealth (190). More details of this study are given in the next section.

Thus, prior to this study relatively little was known about the association between socioeconomic factors and HIV infection in South Africa, although there was some suggestion of a higher risk profile among poorer, less educated and migrant groups. On the basis of literature from this and other countries in sub-Saharan Africa it was hypothesised in this study that higher HIV risk characteristics and HIV infection would be found more commonly among those of lower wealth, migrants and those who had left school. The findings with respect to wealth and migration were therefore somewhat unexpected. As described in the previous section, the power of the study was low to detect anything but a strong underlying association between social factors and HIV infection. Nevertheless, the data presented here are compatible with at least two plausible hypotheses relating to the social patterning of HIV risk in South Africa.

Firstly, it is possible that there has not previously been an association between greater risk of HIV infection and lower household wealth or migration in the study setting, and that this situation has persisted in the study region until at least 2001. Differential HIV awareness among young people may be a product of a wide range of non-socioeconomic forces, for example intellectual capacity, personal choice, parental influence and cultural or religious norms. Similarly, sexual behaviour among adolescents and young adults may be driven by biological forces, cultural norms, parental or guardian supervision and religious beliefs each of which may not be associated with socioeconomic status.

There are differences between the current study and studies from KwaZulu Natal that might partially explain the different findings. The KwaZulu Natal studies focused on long-term migrants from rural districts who were working at high HIV-prevalence destinations such as the Carletonville mines and factories in the port of Richards Bay (186). Migration may have been a more important contributor to HIV risk among this study population. Much of KwaZulu Natal consists of remote rural areas that may be less accessible than the current study area, which sits on the main road to the provincial capital of Polokwane. As well as this, in the KwaZulu Natal research migrants were recruited at destinations with a very high prevalence of HIV infection. Conversely, migration in the current study site might have more often been to local towns and the provincial capital, which may have had a similar HIV prevalence to the study site, as well as to Johannesburg, which has had high HIV prevalence. This combination of factors – more generalised mobility to areas with, on average, a more similar HIV prevalence to the study area – could help explain the lack of an association between migration and prevalent HIV infection among females and the older group of males in the current study. The association between migrancy and HIV seen among younger males was based on small numbers and should be treated with some caution. Nevertheless, this does provide some corroboration for an association between migrancy and HIV risk among young men which would be worthy of further investigation.

Yet given the portrait of migrants in South Africa as being promiscuous and having high risk sexual partnerships, it was perhaps unexpected that there was little evidence of higher numbers of sexual partners among male migrants, although among older, but not younger, females migrants did report more partners. Generally, however, where associations were significant between migrancy status and HIV risk characteristics in the whole group these often suggested lower risk among the migrant group. Among both males and females, migrants were more likely to hold tolerant views about HIV infected individuals and to communicate with parents or household members about HIV. Among those who were sexually active, they did however report fewer partnerships with low frequency of sexual intercourse. Among females, migrants had higher levels of HIV knowledge, were more likely to have had an HIV test and were more likely to report condom use. Other research work has begun to challenge the notion that migrants have higher behavioural risk profiles than non-migrants. Research in South Africa has pointed to the fact that most HIV infections among migrants and their wives in KwaZulu Natal were not contracted from the current partner but from another partner outside the marriage (144). The current study also to some extent corroborates findings from rural Uganda on "travellers", who were reported to hold more accepting views on condoms and to use them more often, and did not have larger numbers of partners or sex with riskier sexual network groups more often than non-travellers (143). It does seem plausible that regular migrants from the current study area, which at the time of the study had relatively poor infrastructure, limited HIV prevention campaigns and relatively low roll-out of life-skills education in schools, may get more access to HIV prevention campaigns in migration destinations. The national LoveLife campaign, for example, whose slogan "Talk about it" relates directly to communication on HIV, may be more active in many migration destinations than in the study area. Migrants may also mix in such destinations with individuals who hold more progressive views and have themselves been exposed to more HIV prevention messages. Similarly, universities and schools away from the site might also have better resourced and more widespread life-skills programmes in line with governments targets. Indeed, among male migrants in the current study, school or college attendance was associated with a greater likelihood of having had an HIV test. No data was collected in the current study on patterns of exposure to HIV prevention campaigns, but it seems plausible given the study findings that that those who migrate are more exposed to such messages. Alternatively, however, traits that lead some individuals to migrate and others not to do so might also be similar traits that lead young people to seek out and digest messages about safe sexual behaviour and other issues affecting their lives. It is possible that temporary migrants in the study population are a self selected group of individuals who *de facto* have lower HIV risk profiles than those who tend to stay at home. Longitudinal data on migration patterns and HIV awareness and other risk characteristics would be useful to understand these dynamics better.

The pattern described above might also help explain why there was no association between wealth and HIV among this South African study population. Previous studies have often interpreted the association between wealth or education and HIV risk characteristics in terms of a pathway by which the more wealthy are more likely to be mobile, and thus have larger sexual networks and more contact with partners from areas with a high background prevalence of HIV infection (153, 165, 171). However, temporary migrancy is a ubiquitous feature of South African life in many rural areas, essential to support the livelihood of most rural households. Temporary migration can be seen as a product of the economic system that affects the whole of areas such as the one in which the current study was conducted. It was therefore perhaps not surprising that temporary migration was not associated with relative household wealth in this study, but was quite common across the whole group. No data was collected in this study on what other factors caused some young people to migrate while others stay at home. Family or social links with households distant from the study site may facilitate migration to areas where jobs are available. Alternatively, those that migrate may be those with psychosocial properties such as increased determination to succeed, while those that stay at home may be more resigned to the low prospects for employment both at home and at migration destinations. Migration was associated in the current study with non-school attendance among the younger group, but overall this accounted for relatively few migrants. At older ages, migrants were as likely to be in school or college as not to be so. Among those not attending school, migration was associated with a greater chance of secure employment, but levels of unemployment were high among both migrants and non-migrants. In general, migration may have been less associated with wealth in this area of South Africa than in areas where previous studies have been conducted, where the association between wealth and HIV risk may indeed be mediated by levels of mobility. So, while previous reports in South Africa have emphasised migration as a key force in promoting risky sexual behaviour, this study challenges that notion. Migration is key to spreading infection at the early stages of an epidemic, but may be less important in mediating epidemic spread as infection spreads to the general population in rural areas. Equally, while migration may be associated with loneliness and lack of community support, thus promoting unsafe sexual practices, so those who remain in rural areas may be poorly served by HIV prevention campaigns.

A second plausible explanation for the lack of an association between wealth and risk of HIV infection in this and other recent studies in South Africa, is that South Africa, as has been reported for some other countries in the region, may be in a transitional stage. Early in the epidemic, HIV may have been concentrated among those from the wealthiest, more educated and most mobile backgrounds. However, over time, HIV may increasingly be being newly transmitted to those of lower wealth. Such a situation could lead to the finding presented in this thesis of little evidence of an association between wealth or migration and risk of HIV infection in 2001; but this may not previously have been the case, and the association may change in the future.

There was some support for this hypothesis in that current school attendance was associated with greater wealth (among females), and also with some lower HIV risk sexual behaviours among both sexes. The pattern identified among this group of young people may be indicative of a trend towards a lower behavioural risk profile among individuals who stay in school longer, become more educated and thus become wealthier. New research might help lend greater insight into the social epidemiology of HIV infection in South

Africa, and thus to a greater understanding of these dynamics more generally. Information on educational attainment has been collected as part of the antenatal clinic surveillance system since its inception, yet it was not possible to identify time-series data on patterning of HIV infection by educational attainment during the literature review conducted for this thesis. Analysis of these data would provide a useful insight into whether social patterning of HIV within South Africa is changing over time. Further monitoring of these data in the future would also be invaluable in this regard. In addition, more nuanced analysis of the social patterning of HIV infection from the recent national surveys of HIV and sexual behaviour conducted in South Africa would be useful. These surveys had larger sample sizes and represented a wider range of wealth than the current study. Additionally, such analyses might help unpick differences between provinces in the patterns identified, as has been hypothesised here. Finally, future national surveys could usefully incorporate data that might be used to provide a more complete picture, such as data on indicators of household wealth. This could help generate greater insight into the mechanism of these associations and their changing nature over the coming years.

10.03.2 The social epidemiology of sexual behaviour and HIV infection among young people

While many previous studies were identified that had previously presented data on the social epidemiology of sexual behaviour and HIV infection, a much smaller number had presented data on young people. Generally, where associations between socioeconomic factors and risk of HIV infection were reported for young people, these related educational attainment to prevalent

HIV infection. Differences were sometimes seen in the association noted among younger and older groups in these analyses (165, 180), and these were generally interpreted as being due to the fact that HIV infections among the young will have been acquired more recently and thus better reflect changing patterns of new infection.

The way in which indicators of education can be interpreted differs between younger and older age groups. Among older groups, greater education may be indicative of increased earning potential or literacy. Among young people, earning benefits have not yet accrued and educational measures may be more indicative of attendance at school and access to current school-based health promotion initiatives. In this study, this difference was made explicit by the use of school attendance as opposed to educational attainment as the socioeconomic exposure investigated. Most previous studies that have adopted this measure have generally done so as a secondary objective and have rarely adjusted analyses for childbirth or for suitably small age groupings. This has made them difficult to interpret. The current study took steps to account for these factors. As hypothesised, the results provided some evidence to suggest that school attendance was associated with lower risk behaviour characteristics and with lower risk of HIV infection among young men. As will be seen in the next section, the mechanisms by which these associations may be mediated may be as much to do with sexual network structure as with knowledge, literacy or empowerment.

This was the first study to present data with respect to the association between migrancy and risk of HIV infection specifically among young people.

Unexpectedly, the results suggested that, if anything, there was a tendency towards a lower HIV risk profile among young temporary migrants. As discussed in the previous section, these results may have been at least partly a consequence of characteristics of the socioeconomic environment of the current study area. However, it is also possible that the results represent a different pattern among young people. Migrancy is less common among young people than among older groups in South Africa (29), will have been for a shorter period of time, and among young people in this study migrancy was as often related to school or college attendance as it was to job seeking or taking up work. It is possible that the negative impacts with respect to migration may take a long time to accrue, perhaps due to loneliness and isolation. Alternatively, it may be specifically labour migration, which is less common among the young, that creates greater risk profiles among migrants. These hypotheses would also be a suitable topic for further research in South Africa and other regions.

Two previously conducted studies provide the most comparable evidence on socioeconomic patterns of sexual behaviour and HIV infection in young people, since both identified the association between measures of socioeconomic position and a number of sexual behaviour characteristics in a manner similar to the current study (147, 190). The results of all three studies are summarised in Table 10.1. As has been suggested by this thesis, the studies may have generated different results due to their use of different methods of wealth estimation that the choice of method of wealth estimation. However, it is difficult to further investigate this hypothesis, and it must be assumed that each chose an appropriate method of relative wealth estimation in their setting.

As in the current study, in both Kisumu, Kenya and KwaZulu Natal, South Africa, condom use among males was not significantly associated with higher socioeconomic position, although in the Kisumu study there was a non-significant trend such that condom use was lower in the lower SES groups. Among females, there was a trend such that condom use was lowest in those of lowest SES in Kisumu, though this was not statistically significant, while in both the current study and KwaZulu Natal, reported condom use was significantly more common among young women of greater wealth.

Table 10.1: Summary of the results of three studies examining the association between household wealth and sexual behaviour characteristics among young people

	Males			Females		
Study	*"Four cities study" (147)*	*"Transitions to adulthood" (190)*	*"IMAGE Study Baseline"*	*"Four cities study" (147)*	*"Transitions to adulthood" (190)*	*"IMAGE Study Baseline"*
Location	*Kisumu*	*Kwazulu Natal*	*Sekhukhuneland*	*Kisumu*	*Kwazulu Natal*	*Sekhukhuneland*
Year	*1996*	*2001*	*2001*	*1996*	*2001*	*2001*
Method to assess household socioeconomic position	Composite marker of education, employment and services	Asset index	Participatory Wealth Ranking	Composite marker of education, employment and services	Asset index	Participatory Wealth Ranking
Age at first sex	No association	No association	No association	Lower age of sexual debut in lowest SES group	Trend towards lower age of sexual debut in lower SES groups	No association
Multiple partners in past year	-	No association	No association	-	Significant association such that wealthiest reported least partners in last year	No association
Secondary abstinence	-	No association	No association	-	No association	No association
Multiple partners during lifetime	No association	-	No association	Borderline significant association such that more partners among higher SES group	-	No association
Age difference with partners	-	No association when first sexual partner considered; significant trend towards wealthier most likely to report wide age gap with recent partner.	No association	No association with age of first spouse	No association with age difference with either first or most recent partner	No association
Condom use	No statistically significant association, although ever condom use lowest in lowest SES group	No association	No clear trend. Some evidence of changing association with age.	No statistically significant association, although ever condom use lowest in lowest SES group	Significant association such that condom use at last sex most often reported by wealthiest	Significant association such that condom use at last sex most often reported by wealthiest
Resource exchange	-	-	No association	-	Strong association such that those of lowest wealth most likely to have "traded sex"	No association
HIV	No association	-	No association	Lower prevalence of HIV infection among higher SES groups	-	No association

With respect to other sexual behaviours, both similarities and differences between the studies were noted. Among males, no association was noted between socioeconomic position and age at first sex in any of the studies. Similarly, no association was noted between high SES and total number of lifetime partners in Kisumu, or having had more than one partner in the last year in Kwazulu Natal, and in the current study no association was noted between household wealth and either of these factors. Both the current and the KwaZulu Natal study found no association between household wealth and secondary abstinence among males. In KwaZulu Natal the largest male to female age gaps were noted among men of the wealthiest group, while in the current study there was little evidence of such a difference. Among females, no association was noted between socioeconomic status and partner ages or secondary abstinence in any of the studies. However, the current study finding of no association between wealth and age of sexual debut among women differed from both previous studies where those of the lowest SES reported the earliest sexual debut. Equally, the current study found no association between household wealth and number of sexual partners (total or in the past year), which was similar to the result seen in Kisumu, but differed from the result of the 'Transitions to Adulthood' study which identified a strong trend such that those of the highest SES were least likely to have reported more than a single sexual partner during the previous year. Finally, in only one study population, young women in Kisumu, was there a trend for lower risk of HIV infection among those of lower socioeconomic position.

In summary, there is growing evidence from the small number of studies conducted in Kenya and South Africa since the mid 1990s that condom use is more commonly reported by young women from households of greater wealth. This association has not been seen as strongly among males, although since fewer males were sexually active in each of the studies the lack of an effect may have been partly a result of decreased statistical power. Age at first sex among males and secondary abstinence among both sexes were not associated in any of the studies. There is less agreement with regard other sexual behaviour markers or HIV infection. However, one common feature of these studies is the relative lack of evidence that young people of either sex from wealthier backgrounds have greater sexual behaviour risk profiles than those of lower SES. Taken together, evidence on sexual behaviour is suggestive of a lower HIV risk profile among those of greater wealth in studies from the mid 1990s onwards, particularly regarding condom use. If patterns of reported sexual behaviour persist it seems likely that new HIV infections will increasingly be transmitted to and among young people of lower wealth. Further research conducted among young people would be useful to support this hypothesis. Such research might also seek to confirm the hypothesis suggested by this study that socioeconomic role is more important than household wealth in determining risk of HIV infection among the young.

10.03.3 Mechanisms underlying the social epidemiology of HIV infection

A secondary aim of this study was to explore the mechanisms by which socioeconomic factors were associated with risk of HIV infection. Review articles and qualitative literature have described a wide variety of theories by which such associations may occur. These include theories linking socioeconomic factors to both higher and lower risk. However, previous empirical studies in this area have generally not analytically investigated the pathways through which socioeconomic factors may be associated with HIV risk. Consequently, there is no consensus in the literature on which pathways are the most important. Nevertheless, there are some trends in interpretation. Generally, where empirical studies have found an association between higher wealth, migration or increased educational attainment and higher HIV risk, this has been taken to represent the same underlying causal pathway. Thus, as noted earlier, explanations of increased HIV risk among those with increased educational attainment have often focused on the association of this factor with increased income and mobility (153, 165, 171). Mobility has generally been seen as a key pathway variable in these relationships, and has been thought to increase HIV risk by increasing contact with sexual partners from populations with a high background prevalence of HIV infection and / or by increasing risky sexual behaviours such as having more partners. More recently, increased recognition of the changing association between educational attainment and HIV infection has promoted the theory that more educated and wealthier groups may be more likely to change their behaviour in response to HIV prevention campaigns, either through increased exposure to or comprehension of their messages, or an increased capacity to act on these (153, 162, 170).

Relative household wealth was generally not associated with HIV risk characteristics for either sex in this study, except that condom use was more common in the wealthier women. One possible reason why there was no association between household wealth and prevalent HIV infection in this study is that relative wealth when compared to neighbours may be less important than absolute wealth, or relative wealth compared more widely, in determining risk behaviours. Those included in the study may have been relatively homogenous with regard absolute levels of wealth. The PWR technique used in this study identified a high level of poverty and demonstrated differences in welfare levels between households in the study sample. Nevertheless, these differences may have been small in absolute terms. Absolute wealth may be most likely to influence risk characteristics if access to resources directly influences contact with sexual partners who have a high probability of infection (for example, commercial sex workers), in situations where sex is traded to provide for basic needs (such as food or schooling), or if there are financial barriers to accessing services that can help reduce the risk of infection such as STI treatment and condoms. It is possible that these factors similarly affected all members of the study population, most of whom were young, unemployed or in school, and would have had limited access to cash resources. Equally, given South Africa's high levels of national inequality coupled with good transport and communication systems, relative wealth compared not just to neighbours but to countrymen with much higher standards of living may effect self esteem and social capital and thus sexual behaviour. This might also have been a phenomenon similarly affecting all young people from the study site.

Where most members of a study sample experience a relatively high level of a putative exposure (in this case, low absolute household wealth), this could be highly important in causing risk of infection, but no, or a very small, gradient may be seen between exposure and outcome. As Geoffrey Rose noted in his work, "Sick Individuals and Sick Populations", if all members of a population smoked 20 cigarettes a day, no amount of epidemiological study would identify smoking as a risk factor for lung cancer (235). In this case other disciplines, such as clinical medicine or cell oncology, might suggest a causative link between smoking and cancer, but observational epidemiological methods could not. Similarly, qualitative study of the narratives of young people can provide insights into the factors driving their sexual behaviours. Many such studies have been conducted in South Africa and these often point to the centrality of poverty in generating behavioural risk profiles (38-40, 106, 236). Qualitative work conducted as part of the IMAGE study has also suggested the importance of resources in determining patterns of relationship formation among women (237). The study might have found an empirical association between wealth and HIV risk if individuals from a wider range of wealth backgrounds had been included. Future analyses of datasets including individuals from a broader spectrum of absolute wealth levels would be useful to better understand the relative importance of relative and absolute wealth in this regard.

The results of this study also suggested that the effects of each of the exposures were independent of the others. Wealth was mostly not associated with outcomes. This persisted either with or without adjustment for schooling or migration and generally there was little evidence that household wealth modified the relationship of these factors with outcomes. Indeed, wealth was not associated with migration among either sex, and was associated with schooling only among females, with this effect being relatively weak since school attendance was relatively common even among the "very poor". The study suggests caution is necessary in future research in interpreting an association of wealth with HIV risk outcomes in terms of schooling or migration patterns, or vice versa. Reference may need to be made to the context of studies in terms of the socioeconomic and HIV risk environments in which they are conducted.

Socioeconomic role variables were more often associated with outcome variables. Independent of household wealth, attendance at school and migration were associated with some lower HIV risk characteristics. The associations between schooling or migration with outcome variables were generally independent of each other. However, there was evidence in a small number of cases of an interdependence of these factors in their association with HIV risk. Most notably, the protective effect of school or college attendance with respect to HIV infection among males was seen among non-migrants only. Thus, among non-migrant males, but not migrants, those who were not attending school were at the greatest risk. This meant that those who were out of school but non-migrant had the highest level of HIV infection. Among females, there was some suggestion that individual sexual behaviour characteristics were also differentially associated with schooling dependent on migrancy status. The reasons for these interactions remain unclear.

The behavioural pathways through which school attendance and migration might be associated with risk of HIV infection were also examined in this thesis. As in previous studies, adjustment for pathway variables generally had little impact on odds ratios (147, 149). However, this may be due to a limitation in cross-sectional data. Migration pathways have been discussed earlier, so the remainder of this section will focus on school attendance.

Attendance at school was associated with a lower number of sexual partners for both men and women, and among women it was associated with other lower risk partnership characteristics such as lower frequency of sex within partnerships, a smaller age difference between partners and increased condom use. Previous literature has tended to emphasise the potential influence of education on awareness of HIV prevention campaigns and / or increased self efficacy to adopt protective behaviours by increasing confidence and hope for the future (162). However, in the current study among both sexes, school attendance was not significantly associated with HIV awareness factors (although there was some suggestion of increased awareness among those attending school among males). An alternative theory suggested by this study is that attendance at school may have an impact on HIV risk behaviours through affecting sexual network structures. The data presented here are compatible with the hypothesis that attendance at school may reduce the size of sexual networks for both sexes, and for females may alter the pool of potential partners such as to encourage the formation of partnerships that have less inherent risk of HIV infection associated with them.

No data were available in this study on the socioeconomic role of sexual partners, but future research may wish to examine the hypothesis that young people attending school, particularly young women, are more likely to form sexual partnerships with other young people attending school. Data on patterns of assortative mixing with respect to economic status or role from southern, eastern or central Africa were not identified in the literature review conducted for this study. However, previous work in other populations has suggested that marital partnerships tend to form with individuals of similar socioeconomic background, although men may be more likely also to form disassortative partnerships in this respect (238). Epidemiological studies have rarely collected information on partner's status, although some studies have found an association between a woman's risk of HIV infection and the educational attainment or income of her partner (146, 156). Future studies might usefully collect information on such issues to further understanding in this area, although the complexities of collecting socioeconomic information on sexual partners should not be underestimated.

An influence of school attendance on sexual network structure could be beneficial for a number of reasons. It may limit the pool of potential sexual partners and thus reduce numbers of partners (or network *node*). Previous work has pointed to the idea that both size of sexual network and position with in it might affect risk of HIV infection (239). Previous work on age differences between sexual partners has tended to emphasis age-imbalance as a potential mediator of the wide difference in HIV prevalence between young men and women at young ages. While there remains debate in this area, it seems clear that wide age differences between partners can be a risk factor for HIV infection among young women and may contribute to their higher prevalence of infection than young men (43, 50, 209, 240), alongside biological mechanisms (241). Socioeconomic inequalities between older men and younger women might also affect the likelihood and circumstances of sexual intercourse occurring, the frequency of sexual intercourse and whether condoms are used (242). This study suggests that school attendance or leaving might also differentially affect the sexual networks of young women and contribute to differential risks, or pathways of risk, among these groups.

In addition to an affect on the characteristics of partners, school attendance might also affect communication characteristics within sexual networks. Adoption of safer sexual behaviours by young people may be facilitated by attendance at school by increasing group bonding leading to group negotiation of positive attitudes towards such behaviours (105, 106). This group negotiation might occur both within gender groups and across them, and school attendance may facilitate the process of negotiation by putting young people in regular contact with each other. School attendance may also encourage the formation of other groups with high levels of social capital that support the adoption of protective behaviours (136). Young people in school may also be more likely to have shared experience and shared reference to HIV prevention campaigns. Conversely, those who leave school may enter adult sexual networks where collective negotiation processes are more fraught, and where older partners with more experience and power dictate the "rules" of sexual engagement.

Few previous studies have examined the impact of school attendance on either HIV or sexual behaviour in sub Saharan African countries. Although the effects noted in this study were modest, they provide some evidence to suggest that attendance at school has the potential to reduce risk of HIV infection, and that this may in part be due to an impact on the sexual networks of students compared to their out of school peers. Future research that builds on these hypotheses would be of great use to further understand the ways in which socioeconomic role can affect risk of HIV infection. In addition, epidemic modelling studies could be used to simulate HIV epidemics with different levels of mixing between partners in different socioeconomic groups to understand how this affects the pattern of prevalent HIV infection. This information could help inform the interpretation of empirical studies that collect such data.

10.03.4 Summary of implications of the current research

The study rejected the hypothesis of increased risk of HIV infection among young people from poorer households. Temporary migrants exhibited a number of lower risk characteristics, although some high risk characteristics were more common in this group. However, the results of the study lent support to the hypothesis that young people attending school would have lower risk profiles. Migration and wealth may be less linked in some areas of South Africa than in other settings, and thus less associated with increased risk of HIV infection. A greater understanding of how the pattern of HIV infection has changed and continues to do so over time would lend great insight into the epidemic in South Africa, but there is some evidence that new infections may be more likely among young people from lower socioeconomic groups. Attendance at school may affect risk of HIV infection among young people by affecting their sexual network structure as well as through psychosocial pathways.

10.04 HIV prevention in sub-Saharan Africa

This study did not evaluate any HIV prevention strategy and has primarily aimed to contribute to better understanding of the social epidemiology of HIV infection among young people. Nevertheless, it is appropriate to conclude this thesis by reflecting briefly on the potential implications of the research findings for strategies to prevent HIV infection among young people in South Africa.

The study findings confirmed that the epidemic of HIV poses a significant public health threat to young people in all socioeconomic groups in these villages in South Africa that had a generalised high level of poverty. The steep rise in HIV prevalence over the age-range studied in this thesis suggested a significant burden of newly acquired HIV infections during this period of life. Therefore, reducing transmission of HIV infection to young people of all social groups should remain an important public health goal in settings such as this. This will require targeting the risky sexual behaviours of young men and women, and additionally the older male partners of young women. The study also lent some support to the idea that young people from some social groups may be at greater risk of HIV infection than others. Thus, young women had a much higher prevalence of HIV infection than men. There was some suggestion that young women from the poorest households were less likely to use condoms than women from wealthier households, and that young men and women not attending school and those who were not migrants had some higher risk characteristics than those who were. Taken together, there was some suggestion that young people from rural areas who leave school early and do not migrate away from the home may be at particularly high risk of infection. It remains unclear from this study, or from available data in South Africa, whether there is a trend towards HIV concentrating in those of lower socioeconomic status, as has been suggested for other settings, though this remains plausible given the available evidence.

The findings of the study support the need to consider at least three responses with respect to HIV prevention in South Africa. Firstly, they suggest that current government-backed HIV prevention strategies must continue to be rolled-out more fully in areas such as the current study site; secondly, they suggest a need to implement these strategies more effectively; finally, they provide some support for calls to complement existing strategies with new, more structural, approaches to HIV prevention.

HIV prevention efforts in South Africa are guided by a government endorsed strategic plan (243). This strategic plan emphasises a number of strategies of particular relevance to the prevention of HIV among young people, including making voluntary counselling and testing services widely available and accessible to young people, promoting safe sexual behaviour through school and multimedia programmes, and improving the management and control of STIs. However, in this study few young people of either sex had ever had an HIV test, there were important gaps in HIV knowledge and unsafe sexual behaviour practices remained common. VCT services had not been effectively rolled out in many areas of South Africa by 2001 (244), including in the study site. Efforts to make clinics more adolescent friendly, and thus facilitate condom distribution and improved health seeking behaviour for those with STIs, were also at nascent stages at this time. Similarly, while legislation has supported the introduction of a life-skills curriculum in schools, this program has been slow to effectively roll-out in many provinces of South Africa, notably Limpopo province (245). Challenges to implementing AIDS policy in South Africa have been many and include the legacy of the apartheid system, which left a health and education system with constrained infrastructure after many years of neglect, as well as a perceived lack of will on the part of the government (198). Redoubled efforts to ensure the swift and equitable implementation of government policy despite these barriers are needed.

The current study also suggests that current programmes may need to be refined to maximise their impact. For example, HIV prevention campaigns have tended to emphasise the ABC approach to HIV prevention ("Abstain, Be faithful, Condomise"). However, the importance of age-mixing between younger females and older males has received less attention. This is despite this practice being widespread in South Africa, its importance in HIV transmission being widely recognised and calls for an increased focus on this in HIV prevention campaigns (209). HIV prevention programmes may also need to make increased efforts to reach certain groups. The lower HIV awareness of non-migrants in the current study suggests that campaigns may not be reaching or effectively getting their message across to young people from rural areas. While multimedia campaigns have achieved wide outreach in South Africa, this has not been universal, and such campaigns have been most noticeable in relatively well-resourced and accessible areas. Questions have also been raised about the accessibility of the national LoveLife campaign, which has been accused of being deliberately obscure in its messages and irrelevant to many young South Africans, particularly those from rural and deprived areas (246). The current study suggests that programmes that reach young people who are no longer in school, who may also be particularly at risk, may need to be strengthened. It seems likely that at least some young women were not attending school because they had already fallen pregnant and/or had a child, strengthening calls for HIV prevention strategies in both schools and communities to target young people before they become sexually active.

Finally, many authors have called for structural interventions to be deployed to complement existing HIV prevention strategies. While definitions vary, such interventions recognise the importance of social, economic, legal and cultural forces in determining vulnerability to HIV and target these distal determinants of risk rather than more proximate risk factors (126, 247). In settings with generalised HIV epidemics, poverty, gender inequalities in opportunity and migration have been particularly highlighted as key forces driving the spread of HIV (128). Indeed, some authors have suggested that poverty reduction is perhaps the only sustainable solution to the continued spread of HIV in low-resource settings (248).

This study suggested that strategies that improve the relative wealth of households within the study site may be relatively ineffective in reducing transmission of HIV if they have no knock-on effects, under the important caveat that more needs to be known about whether social patterning of new infections is changing over time. Structural strategies have also been suggested that may be effective in reducing transmission of HIV, but about which this study does not directly provide any information, such as those that promote gender equality of opportunity, affect the absolute level of wealth or social cohesion across the community or impact societal levels of wealth inequality. These interventions may of course have important other benefits that support their introduction in their own right.

However, the results of this study do provide support for some potential structural approaches to HIV prevention. These would include strategies that increase levels of school or college attendance and those that improve the coverage and quality of HIV prevention activities in rural areas.

A number of policy initiatives have been implemented since 1994 to improve education in South Africa. This has included a legal requirement that no child can be excluded from school due to an inability to meet school fees (249). To support this legal requirement, legislation allows school governing boards to set their own fees, but poor households may apply for partial or full exemption from these fees. However, problems do remain for the poor in ensuring their children can attend school. These include that few households have applied for school fee exemptions; hidden costs of schooling (such as contributions to books, school upkeep, excursions etc) may be high and are not covered by exemptions; and, indirect costs of schooling (such as uniforms or transport) may be even higher (250). Consequently, many children are absent from school some of the time, and few complete the secondary school qualification (20). Suggested modifications to the current system that may improve school attendance rates, and have the potential to have a knock-on effect on HIV risk behaviours, include more input from the Department of Education in helping schools set fees and incorporating hidden costs within these fees, better implementation of the school fee exemption programme and state coverage of indirect costs of schooling through, for example, providing transport for those in remote rural areas (250). A more radical suggestion may be to offer financial incentives to poor households to keep their children in school or college. In south and central America, direct cash transfers have been used to incentives poor households to access basic health services such as immunisations (251, 252). This would of course have important implications in terms of funding and management to be effectively implemented.

Community based interventions might also be used to improve school enrolment rates. In particular, evaluations of microfinance initiatives for poor households suggest that increased school enrolment is a key benefit of participation in these programmes (253-256). Finally, initiatives that improve the quality of teaching in schools and maximise the number of students who successfully complete matriculation are urgently needed, and may have knock-on effects on HIV prevention.

As suggested earlier, promoting better HIV prevention campaigns that reach those who remain in rural areas may be achieved through targeting and outreach to these areas. More structural interventions or policies to facilitate this could include incentivising NGOs and other organisations to conduct programmes in such areas through grant schemes. Promoting the use of community development approaches to mobilise communities and maximise the effectiveness of pre-existing resources and knowledge may also be a useful strategy to improve the effectiveness of information and education campaigns in such areas. Finally, stimulating general economic development to increase the accessibility of and levels of economic activity in rural areas may improve the quality of services delivered in these regions.

In conclusion, structural interventions for HIV prevention have received much attention of late, and may be an important strategy in long term efforts to halt the spread of HIV infection. However, there have been few attempts to rigorously evaluate such programmes. What seems clear is that such research, alongside rapid expansion and monitoring and evaluation of existing efforts, remain urgent public health priorities in countries affected by generalised HIV epidemics.

10.05 Conclusion

The epidemic of HIV remains a serious public health threat to the health of young people from all social groups in sub-Saharan Africa. This study was conducted in a former bantustan in rural South Africa with high levels of poverty and unemployment and that was relatively homogenous with regard absolute wealth. It found little evidence of an association between relative household wealth and current risky sexual behaviour or HIV infection, except that females from the poorest households were least likely to use condoms. However, there was evidence that some characteristics associated with lower risk of HIV infection were more common among young people attending school and temporary migrants. It will be important to monitor patterns of risky sexual behaviour and HIV infection in South Africa over time, to see whether risk concentrates among those of lower wealth, education and mobility. Since poorer people may have least access to HIV/AIDS treatment programmes as they expand within sub-Saharan Africa, it is possible that an increased emphasis on HIV prevention campaigns that are effective among this group may anyway be necessary. Rapid expansion of existing efforts, alongside evaluation of both targeted and structural interventions for HIV prevention among young people are clear priorities for action.

References

1. Shucksmith J, Hendry LB. Health issues and adolescents: growing up, speaking out. London New York: Routledge; 1998.

2. Viner R, Booy R. Epidemiology of health and illness. British Medical Journal 2005;330(7488):411-4.

3. Viner R, Macfarlane A. Health promotion. British Medical Journal 2005;330(7490):527-9.

4. Crockett LJ, Petersen AC. Adolescent development: Health risks and opportunities for health promotion. In: Millstein SG, Petersen AC, Nightingale EO, editors. Promoting the health of adolescents. New York: Oxford University Press; 1993.

5. Christie D, Viner R. Adolescent development. British Medical Journal 2005;330(7486):301-4.

6. Delius P. A Lion amongst the Cattle. Portsmouth: Heinemann; 1996.

7. Sparks A. Beyond the miracle. Inside the New South Africa. Johannesburg & Cape Town: Jonathan Ball Publishers; 2003.

8. Moll P. Discrimination is declining in South Africa but inequality is not. Journal of Economics and Econometrics 2000;24(3):91.

9. Rose D, Charlton KE. Prevalence of household food poverty in south Africa: results from a large, nationally representative survey. Public Health Nutrition 2003;5(3):383-389.

10. May J, Carter M, Posel D. The composition and persistence of poverty in rural South Africa:An entitlements approach. Johannesburg: Land and Agriculture Policy Centre; 1995.

11. May J. Experiences and perceptions of poverty in South Africa. Durban: Praxis Publishing; 1998.

12. Posel D. How do households work? Migration, the household and remittance behaviour in South Africa. Social Dynamics-a Journal of the Centre for African Studies University of Cape Town 2001;27(1):165-189.

13. Posel D, Casale D. What has been happening to internal labour migration in South Africa, 1993-1999? South African Journal of Economics 2003;71(3):455-479.

14. Labour Force Survey. Pretoria: Statistics South Africa; 2002.

15. Kingdon G, Knight J. What have we learnt about unemployment from microdatasets in South Africa? Social Dynamics-a Journal of the Centre for African Studies University of Cape Town 2001;27(1):79-95.

16. Kingdon GG, Knight J. Unemployment in South Africa: The nature of the beast. World Development 2004;32(3):391-408.

17. Carter M, May J. Poverty, livelihood and class in rural South Africa. World Development 1999;27(1):1-20.

18. Anderson KG, Case A, Lam D. Causes and consequences of schooling outcomes in South Africa: Evidence from survey data. Social Dynamics-a Journal of the Centre for African Studies University of Cape Town 2001;27(1):37-59.

19. Rutenberg N, Kehus-Alons C, Brown L, Macintyre K, Dallimore A, Kaufman C. Transitions to adulthood in the context of AIDS in South Africa. New York: Population Council; 2001.

20. Hallman K, Grant M. Poverty, educational attainment and livelihoods: How well do young people fare in KwaZulu Natal, South Africa. Washington, D.C.: Population Council; 2004.

21. Braemher S, Kimmia Z, Greenstein R, Morale R, Sektloadi K. Youth 2000: a study of youth in South Africa. Braamfontein: Community Agency for Social Enquiry; 2000.

22. Lestrade-Jefferis J. The labour market. In: Udjo EO, editor. The People of South Africa Population Census 1996. Pretoria: Statistics South Africa; 2000.

23. The South African labour market: selected time-based social and international comparisons. Pretoria: Statistics South Africa; 2002.

24. Ramphele M. Steering by the stars. Being young in South Africa. Cape Town: Tafelberg; 2002.

25. Schoeman NJ, Blignaut JN. Socio-economic environment and labour absorption in South Africa. South African Journal of Economics 1998;66(3):299-318.

26. Network YD. Southern African perspectives on youth employment; 2004.

27. Lurie M. Migration and AIDS in southern Africa: a review. South African Journal of Science 2000;96:343-347.

28. Hosegood V. An alternative approach to defining households in Southern Africa. In: Van de Walle E, editor. African Households. Illinois: Illinois Press; 2003.

29. Kahn K, Collinson M, Tollman S, Wolff B, Garenne M, Clark SJ. Health consequences of migration: Evidence from South Africa's rural northeast. In: African Migration in Contemporary Perspective; 2993; Johannesburg; 2993.

30. Madhavan S, Townsend N, Tollman S, Garenne M. Children's residence patterns and well being in rural South Africa. In; 2000 2000; Los Angeles; 2000.

31. Collinson M, Wittenburg M. Labour force dynamics in a rural part of South Africa: the Agincourt sub-district of the Northern Province, 1992 – 2000. Johannesburg: University of the Witwatersrand; 2001.

32. Department of Health. National HIV and syphillis antenatal sero-prevalence survey in South Africa 2003. Pretoria: Department of Health South Africa; 2003.

33. Manzini N. Sexual initiation and childbearing among adolescent girls in KwaZulu Natal, South Africa. Reprod Health Matters 2001;9(17):44-52.

34. Jewkes R, Vundule C, Maforah F, Jordaan E. Relationship dynamics and teenage pregnancy in South Africa. Social Science and Medicine 2001;52(5):733-44.

35. Kelly K, Parker W. Communities of practice. Contextual mediators of youth response to HIV/AIDS. Sentinel site monitoring and evaluation project.; 2000.

36. DOH. South Africa Demographic and Health Survey - 1998. Pretoria: Department of Health, South Africa; 1998.

37. Delius P, Glaser C. Sexual socialisation in historical perspective. In: AIDS in Context, Johannesburg; 2001.

38. Collins T, Stadler J. Love, Passion and Play:Sexual meaning among youth in the Northern Province of South Africa. In: AIDS in Context, Johannesburg; 2001.

39. Varga CA. Sexual decision-making and negotiation in the midst of AIDS: youth in Kwazulu Natal, South Africa. Health Transition Review (supplement) 1997;7:45-67.

40. Kaufman C, Stavrou SE. "Bus fare, please": The economics of sex and gifts among adolescents in urban South Africa. New York: Population Council; 2002.

41. Slaymaker E, Zaba B. Measurement of condom use as a risk factor for HIV infection. Reprod Health Matters 2003;11(22):174-84.

42. Auvert B, Ballard R, Campbell C, Carael M, Carton M, Fehler G, et al. HIV infection among youth in a South African mining town is associated with herpes simplex virus-2 seropositivity and sexual behaviour. AIDS 2001;15(7):885-898.

43. Macphail C, Williams B, Campbell C. Relative risk of HIV infection among young men and women in a South African township. International Journal of STD and AIDS 2002;13:331-342.

44. Flisher AJ, Reddy P, Muller M, Lombard C. Sexual behaviour of Cape Town high-school students. South African Medical Journal 2003;93(7):537-41.

45. Shisana O, Simbayi L. South African National HIV Prevalence, Behavioral Risks and Mass Media. Household Survey 2002. Cape Town: Human Sciences Research Council; 2002.

46. Pettifor AE, Rees HV, Steffenson A, Hlongwa-Madikizela L, MacPhail C, Vermaak K, et al. HIV and sexual behaviour among young South Africans: a national survey of 15-24 year olds. Johannesburg: Reproductive Health Research Unit, University of the Witwatersrand; 2004.

47. Leclerc-Madlala S. Infect one, infect all:Zulu youth response to the AIDS epidemic in South Africa. Medical Anthropology 1997;17:363-380.

48. Buga GA, Amoko DH, Ncayiyana DJ. Sexual behaviour, contraceptive practice and reproductive health among school adolescents in rural Transkei. South African Medical Journal 1996;86(5):523-7.

49. Williams BG, Taljaard D, Campbell CM, Gouws E, Ndhlovu L, van Dam J, et al. Changing patterns of knowledge, reported behaviour and sexually transmitted infections in a South African gold mining community. AIDS 2003;17:2099-2107.

50. Luke N. Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa. Studies in Family Planning 2003;34(2):67-86.

51. MacPhail C, Campbell C. 'I think condoms are good but, aaai,I hate those things':condom use among adolescents and young people in a Southern African township. Social Science and Medicine 2001;52:1613-1627.

52. Wood K, Maforah F, Jewkes R. "He forced me to love him": putting violence on adolescent sexual health agendas. Social Science and Medicine 1998;47(2):233-42.

53. Stadler J. Rumor, gossip and blame: implications for HIV/AIDS prevention in the South African lowveld. AIDS Education and Prevention 2003;15(4):357-68.

54. Abdool Karim SS, Abdool Karim Q, Preston-Whyte E, Sankar N. Reasons for lack of condom use among high school students. South African Medical Journal 1992;82:107-11.

55. Gilmour E, Karim SS, Fourie HJ. Availability of condoms in urban and rural areas of KwaZulu-Natal, South Africa. Sexually Transmitted Diseases 2000;27(6):353-7.

56. Camlin CS, Chimbwete CE. Does knowing someone with AIDS affect condom use? An analysis from South Africa. AIDS Education and Prevention 2003;15(3):231-44.

57. Jewkes R, Penn-Kekana L, Levin J, Ratsaka M, Schrieber M. Prevalence of emotional, physical and sexual abuse of women in three South African provinces. South African Medical Journal 2001;91(5):421-8.

58. Jewkes R, Abrahams N. The epidemiology of rape and sexual coercion in South Africa: an overview. Social Science and Medicine 2002;55(7):1231-44.

59. Jewkes R, Levin J, Mbananga N, Bradshaw D. Rape of girls in South Africa. Lancet 2002;359:319-320.

60. Andersson N, Ho-Foster A, Matthis J, Marokoane N, Mashiane V, Mhatre S, et al. National cross sectional study of views on sexual violence and risk of HIV infection and AIDS among South African school pupils. British Medical Journal 2004;329(7472):952.

61. Abrahams N, Jewkes R, Hoffman M, Laubsher R. Sexual violence against intimate partners in Cape Town: prevalence and risk factors reported by men. Bulletin of the World Health Organisation 2004;82(5):330-7.

62. Garcia-Moreno C, Watts C. Violence against women: its importance for HIV/AIDS. AIDS 2000;14 Suppl 3:S253-65.

63. Dunkle KL, Jewkes RK, Brown HC, Yoshihama M, Gray GE, McIntyre JA, et al. Prevalence and patterns of gender-based violence and revictimization among women attending antenatal clinics in Soweto, South Africa. American Journal of Epidemiology 2004;160(3):230-9.

64. Mhlongo S, Fiala C, De Harven E, Rasnick D, Stewart GT. HIV a sexually transmitted disease? An analysis of the latest antenatal screening for HIV and syphilis from South Africa. International Journal of STD and AIDS 2003;14(8):574-6.

65. Schmid GP, Buve A, Mugyenyi P, Garnett GP, Hayes RJ, Williams BG, et al. Transmission of HIV-1 infection in sub-Saharan Africa and effect of elimination of unsafe injections. Lancet 2004;363(9407):482-8.

66. Weiss HA, Buve A, Robinson NJ, Van Dyck E, Kahindo M, Anagonou A, et al. The epidemiology of HSV2 infection and its association with HIV infection in four African populations. AIDS 2001;15 (suppl 4):S97-S108.

67. Weiss HA, Quigley MA, Hayes RJ. Male circumcision and risk of HIV infection in sub-Saharan Africa: a systematic review and meta-analysis. AIDS 2000;14(15):2361-70.

68. Rain-Taljaard RC, Lagarde E, Taljaard DJ, Campbell C, MacPhail C, Williams B, et al. Potential for an intervention based on male circumcision in a South African town with high levels of HIV infection. AIDS Care 2003;15(3):315-27.

69. Mayatula V, Mavundla TR. A review on male circumcision procedures among South African blacks. Curationis 1997;20(3):16-20.

70. Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, Puren A. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial. PLoS Med 2005;2(11):e298.

71. Wilkinson D, Abdool Karim SS, Williams B, Gouws E. High HIV incidence and prevalence among young women in rural South Africa: developing a cohort for intervention trials. Journal of Acquired Immune Deficiency Syndromes 2000;23(5):405-9.

72. Gouws E, Williams BG, Sheppard HW, Enge B, Karim SA. High incidence of HIV-1 in South Africa using a standardized algorithm for recent HIV seroconversion. Journal of Acquired Immune Deficiency Syndromes 2002;29(5):531-5.

73. Dorrington R, Bourne D, Bradshaw D, Laubscher R, Timaeus IM. The impact of HIV/AIDS on adult mortality in South Africa. South Africa: Medical Research Council; 2001.

74. Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. Epidemiologic Reviews 1988;10:87-121.

75. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies and guidelines. Annual Review of Public Health 1997;18:341-378.

76. Wardle J, Robb K, Johnson F. Assessing socioeconomic status in adolescents: the validity of a home affluence scale. Journal of Epidemiology and Community Health 2002;56:595-599.

77. Martin WG, Beittel M. The hidden abode of reproduction:conceptualizing households in Southern Africa. Development and Change 2002;18:215-234.

78. Smit W. The rural linkages of urban households in Durban, South Africa. Environment and Urbanization 1998;10(1):77-87.

79. Chambers R, Conway GR. Sustainable Rural Livelihoods : Practical concepts for the 21st century. Brighton: Institute of Development Studies; 1992.

80. Department for International Development. Sustainable Livelihoods Guidance Sheets. www.dfid.gov.uk: Department for International Development; 2002.

81. Mcdowell C, de Haan A. Migration and sustainable livelihoods:A critical review of the literature. Brighton: Institute of Development Studies; 1998.

82. Myer L, Ehrlich RI, Susser ES. Social epidemiology in South Africa. Epidemiology Reviews 2004;26:112-23.

83. Wojcicki JM. 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(1):1-36.

84. Ravallion M. Issues in measuring and modelling poverty. The Economic Journal 1996;106(438):1328-43.

85. Ravallion M. Poverty Comparisons : a guide to concepts and methods. Washington DC: The World Bank; 1992.

86. Falkingham J, Namazie C. Measuring health and poverty: a review of approaches to identifying the poor. London: DFID; 2002.

87. Morris SS, Carletto C, Hoddinott J, Christiaensen LJM. Validity of rapid estimates of household wealth and income for health surveys in rural Africa. Journal of Epidemiology and Community Health 2000;54:381-387.

88. Schellenberg J-A, Victora CG, Mushi A, de Savigny D, Schellenberg D, Mshinda H, et al. Inequities among the very poor: health care for children in rural southern Tanzania. Lancet 2003;361:561-566.

89. Sahn DE, Stifel D. Exploring alternative measures of welfare in the absence of expenditure data. Review of income and wealth 2003;49(4):463-489.

90. Booysen Fle R. Using demographic and health surveys to measure poverty- an application to South Africa. Journal of Economics and Econometrics 2002;26(3):53-70.

91. Durkin MS, Islam S, Hasan ZM, Zaman SS. Measures of socioeconomic status for child heath research: comparative results from Bangladesh and Pakistan. Social Science and Medicine 1994;38(9):1289-1297.

92. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data - or tears: an application to educational enrolments in states of India. Demography 2001;38(1):115-132.

93. Henry C, Sharma M, Lapenu C, Zeller M. Assessing the relative poverty of microfinance clients. A CGAP operational tool. Washington D.C.: International Food Policy Research Institute; 2000.

94. Chambers R. The origins and practice of participatory rural appraisal. World Development 1994;27(7):953-969.

95. McGee R. Analysis of participatory povety assessment (PPA) and household survey findings on poverty trends in Uganda. Brighton: Institute for Development Studies; 2000.

96. Adams AM, Evans TG, Mohammed R, Farnsworth J. Socioeconomic stratification by wealth ranking: Is it valid? World Development 1997;25(7):1165-1172.

97. Scoones I. Investigating difference: Applications of weath ranking and household survey approaches among farming households in Southern Zimbabwe. Development and Change 1995;26:67-88.

98. Temu AE, Due JM. Participatory appraisal approaches versus sample survey data collection: a case of smallholder farmers well-being ranking in Njombo District, Tanzania. Journal of African Economies 2000.

99. Azjen I, Fishbein M. Understanding attitudes and predicting social behaviour. New Jersey: Prentice Hall; 1980.

100. Rosenstock I. Historical Origins of the Health Belief Model; 1974.

101. Bandura A. Social learning theory. New Jersey: Prentice Hall; 1977.

102. Fishbein M. The role of theory in HIV prevention. AIDS Care 2000;12(3):273-278.

103. Prochaska JO, DiClemente CC. Transtheoretical therapy toward a more integrative model of change. Psychotherapy: Theory, Research and Practice 1982;19(3):276-87.

104. Catania J, Kegeles S, Coates T. Towards an understanding of risk behaviour:an AIDS risk reduction model. Health Education Quarterly 1990;17(1):53-72.

105. Campbell C, Macphail C. Peer education, gender and the development of critical consciousness: participatory HIV prevention by South African youth. Social Science and Medicine 2002;55:331-345.

106. Campbell C. "Letting them die". Why HIV/AIDS prevention programmes fail. Oxford: James Currey; 2003.

107. Campbell C, Williams B, Gilgen D. Is social capital a useful conceptual tool for exploring community level influences on HIV infection? An exploratory case study from South Africa. AIDS Care 2002;14(1):41-54.

108. Gisselquist D, Rothenberg R, Potterat J, Drucker E. HIV infections in sub-Saharan Africa not explained by sexual or vertical transmission. International Journal of STD & AIDS 2002;13(10):657-66.

109. Anderson RM, Grassly N, Garnett GP. Success in HIV control:fact or fiction? In; 2000 2000; 2000.

110. Bwayo J, Plummer F, Omari M, Mutere A, Moses S, Ndinya-Achola J, et al. Human immunodeficiency virus infection in long-distance truck drivers in east Africa. Archives of Internal Medicine 1994;154:1391-1396.

111. Anderson RM, May RM, Boily MC, Garnett GP, Rowley JT. The spread of HIV-1 in Africa: sexual contact patterns and the predicted demographic impact of AIDS. Nature 1991;352:581-589.

112. Gregson S, Nyamukapa CA, Garnett GP, Mason PR, Zhuwau T, Carael M, et al. Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. Lancet 2002;359:1896-1903.

113. Halperin D. Is poverty the root cause of (southern) African AIDS. In: AIDS in Context, Johannesburg; 2001.

114. Wawer MJ, Gray R, Serwadda D, Namukwaya Z, Makumbi F, Sewankambo N, et al. Declines in HIV prevalence in Uganda: not as simple as ABC (Abstract no: 27LB). In: 12th Conference on Retroviruses and Opportunistic Infections; 2005; Boston; 2005.

115. Galvin SR, Cohen MS. The role of sexually transmitted diseases in HIV transmission. Nat Rev Microbiol 2004;2(1):33-42.

116. Grosskurth H, Mosha F, Todd J, Mwijarubi E, Klokke A, Senkoro K, et al. Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: Randomised controlled trial. Lancet 1995;346(8974):530-536.

117. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Paxton L, Berkley S, et al. A randomized, community trial of intensive sexually transmitted disease control for AIDS prevention, Rakai, Uganda. AIDS 1998;12(10):1211-1225.

118. Kamali A, Kinsman J, Nalweyiso N, Mitchell K, Kanyesigye E, Kengeya-Kayondo J, et al. A community randomized controlled trial to investigate impact of improved STD management and behavioural interventions on HIV incidence in rural Masaka, Uganda: trial design, methods and baseline findings. Tropical Medicine and International Health 2002;7(12):1053-1063.

119. Kelly R, Kiwanuka N, Wawer MJ, Serwadda D, Sewankambo NK, Wabwire-Mangen F, et al. Age of male circumcision and risk of prevalent HIV infection in rural Uganda. AIDS 1999;13(3):399-405.

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

121. Watts CH, May RM. The influence of concurrent partnerships on the dynamics of HIV/AIDS. Math Biosci 1992;108(1):89-104.

122. Campbell C. Migrancy, masculine identities and AIDS: the psychosocial context of HIV transmission on the South African gold mines. Social Science and Medicine 1997;45(2):273-81.

123. Barnett A, Whiteside A. AIDS in the 21st Century: Palgrave Macmillan; 2002.

124. Decosas J. HIV and Development. AIDS 1996;10 (suppl 3):S69-S74.

125. Over M. The effects of societal variables on urban rates of HIV infection in developing countries:An exploratory analysis. In: Confronting AIDS:Public Priorities in a Global Epidemic. New York: Oxford University Press; 1998.

126. Sweat MD, Denison J. Reducing HIV incidence in developing countries with social and structural interventions. AIDS 1995;9 (suppl A):s251-s257.

127. Sumartojo E. Structural factors in HIV prevention: concepts, examples, and implications for research. AIDS 2000;14 (suppl 1):S3-S10.

128. Parker RG, Easton D, Klein CH. Structural barriers and facilitators in HIV prevention: A review of international research. AIDS 2000;14 (suppl 1):1.

129. Over M, Piot P. HIV Infection and Sexually Transmitted Diseases. In: Jamison DT, Mosley WH, Mensham AR, Bobadilla JL, editors. Disease Control Priorities in Developing Countries: Oxford University Press; 1993. p. 455-527.

130. Boerma JT, Weir SS. Integrating demographic and epidemiological approaches to research on HIV/AIDS: the proximate-determinants framework. Journal of Infectious Diseases 2005;191 Suppl 1:S61-7.

131. Eaton L, Flisher AJ, Aaro LE. Unsafe sexual behaviour in South African youth. Social Science and Medicine 2003;56(1):149-65.

132. Marmot M. Self esteem and health. British Medical Journal 2003;327(7415):574-5.

133. Gilbert L, Walker L. Treading the path of least resistance: HIV/AIDS and social inequalities a South African case study. Social Science and Medicine 2002;54(7):1093-110.

134. Farmer P. Infections and inequalities: the modern plagues. Berkley: University of California Press; 1999.

135. Gregson S, Waddell H, Chandiwana SK. School education and HIV control in sub-Saharan Africa: from discord to harmony? Journal of International Development 2001;13(4):467.

136. Gregson S, Terceira N, Mushati P, Nyamukapa C, Campbell C. Community group participation: can it help young women to avoid HIV? An exploratory study of social capital and school education in rural Zimbabwe. Social Science and Medicine 2004;58(11):2119-32.

137. Blanc AK. The Relationship Between Sexual Behavior and Level of Education in Developing Countries; 2001.

138. Gage AJ, Meekers D. Sexual activity before marriage in sub-Saharan Africa. Social Biology 1994;41(1-2):44-60.

139. Nyanzi B, Nyanzi S, Wolff B, Whitworth J. Money, men and markets: Economic and sexual empowerment of market women in southwestern Uganda. Culture, Health & Sexuality 2005;7(1):13-26.

140. Chirwa WC. Migrant labour, sexual networking and multi-partnered sex in Malawi. Health Transition Review 1997;7 (Suppl 3):5-15.

141. Jochelson K, Mothibeli M, Leger J-P. Human immunodedeficiency virus and migrant labour in South Africa. International Journal of Health Services 1991;21(1):157-173.

142. Lydie N, Robinson NJ, Ferry B, Akam E, De Loenzien M, Zekeng L, et al. Adolescent sexuality and the HIV epidemic in Yaounde, Cameroon. Journal of Biosocial Science 2004;36(5):597-616.

143. Morris M, Wawer MJ, Makumbi F, Zavisca JR, Sewankambo N. Condom acceptance is higher among travelers in Uganda. AIDS 2000;14(6):733-41.

144. Lurie MN, Williams BG, Zuma K, Mkaya-Mwamburi D, Garnett GP, Sweat MD, et al. Who infects whom? HIV-1 concordance and discordance among migrant and non-migrant couples in South Africa. AIDS 2003;17:2245-2252.

145. Anarfi JK, Appiah EN, Awusabo-Asare K. Livelihood and the risk of HIV/AIDS infection in Ghana: the case of female itinerant traders. Health Transition Review 1997;7 Suppl:225-42.

146. Allen S, Lindan C, Serufilira A, Van de Perre P, Rundle AC, Nsengumuremyi F, et al. Human immunodeficiency virus infection in urban Rwanda. Demographic and behavioral correlates in a representative sample of childbearing women. Journal of the American Medical Association 1991;266(12):1657-1663.

147. Hargreaves JR. Socioeconomic status and risk of HIV infection in an urban population in Kenya. Tropical Medicine and International Health 2002;7(9):793-802.

148. Barongo LR, Borgdorff MW, Mosha FF, Nicoll A, Grosskurth H, Senkoro KP, et al. The epidemiology of HIV-1 infection in urban areas, roadside settlements and rural villages in Mwanza Region, Tanzania. AIDS 1992;6(12):1521-1528.

149. Glynn JR, Carael M, Buve A, Anagonou S, Zekeng L, Kahindo M, et al. Does increased general schooling protect against HIV infection? A study in four African cities. Tropical Medicine and International Health 2004;9(1):4-14.

150. Gregson S, Mason PR, Garnett GP, Zhuwau T, Nyamukapa CA, Anderson RM, et al. A rural HIV epidemic in Zimbabwe? Findings from a population-based survey. International Journal of STD and AIDS 2001;12(3):189-196.

151. Serwadda D, Wawer MJ, Musgrave SD, Sewankambo NK, Kaplan JE, Gray RH. HIV risk factors in three geographic strata of rural Rakai District, Uganda. AIDS 1992;6(9):983-989.

152. Lagarde E, Carael M, Glynn JR, Kanhonou L, Abega SC, Kahindo M, et al. Educational level is associated with condom use within non-spousal partnerships in four cities of sub-Saharan Africa. AIDS 2001;15(11):1399-408.

153. Hargreaves JR, Glynn JR. Educational attainment and HIV-1 infection in developing countries: a systematic review. Tropical Medicine and International Health 2002;7(6):489-98.

154. Glynn JR, Ponnighaus J, Crampin AC, Sibande F, Sichali L, Nkhosa P, et al. The development of the HIV epidemic in Karonga District, Malawi. AIDS 2001;15(15):2025-9.

155. Crampin AC, Glynn JR, Ngwira BM, Mwaungulu FD, Ponnighaus JM, Warndorff DK, et al. Trends and measurement of HIV prevalence in northern Malawi. AIDS 2003;17(12):1817-25.

156. Dallabetta GA, Miotti PG, Chiphangwi JD, Saah AJ, Liomba G, Odaka N, et al. High socioeconomic status is a risk factor for human immunodeficiency virus type 1 (HIV-1) infection but not for sexually transmitted diseases in women in Malawi: Implications for HIV-1 control. Journal of Infectious Diseases 1993;167(1):36-42.

157. Chao A, Bulterys M, Musanganire F, Habimana P, Nawrocki P, Taylor E, et al. Risk factors associated with prevalent HIV-1 infection among pregnant women in Rwanda. National University of Rwanda-Johns Hopkins University AIDS Research Team. International Journal of Epidemiology 1994;23(2):371-80.

158. Kwesigabo G, Killewo J, Godoy C, Urassa W, Mbena E, Mhalu F, et al. Decline in the prevalence of HIV-1 infection in young women in the Kagera region of Tanzania. Journal of Acquired Immune Deficiency Syndromes 1998;17(3):262-268.

159. Carael M. Sexual Behaviour. In: Cleland J, Ferry B, editors. Sexual behaviour and AIDS in the developing World. London: Taylor&Francis; 1995.

160. Deheneffe J-C, Carael M, Noumbissi A. Socioeconomic determinants of sexual behavior and condom use. In: Confronting AIDS:Public priorities in a global epidemic. New York: Oxford University Press; 1998.

161. de Walque D. How does the impact of an HIV/AIDS information campaign vary with educational attainment? Evidence from rural Uganda. Chicago: University of Chicago; 2002.

162. de Walque D, Nakiyingi-Miiro JS, Busingye J, Whitworth JA. Changing association between schooling levels and HIV-1 infection over 11 years in a rural population cohort in south-west Uganda. Tropical Medicine and International Health 2005;10(10):993-1001.

163. Wannan GJ, Cutting WA, Fischer PR. HIV seroprevalence amongst pregnant women in northeastern Zaire. International Journal of STD and AIDS 1997;8(5):317-9.

164. Senkoro K, Boerma JT, Klokke A, Ng'weshemi JZL, Muro AS, Gabone R, et al. HIV incidence and HIV-associated mortality in a cohort of factory workers and their spouses in Tanzania, 1991 through 1996. Journal of Acquired Immune Deficiency Syndromes 2001;23:194-202.

165. Smith J, Nalagoda F, Wawer MJ, Serwadda D, Sewankambo N, Konde-Lule J, et al. Education attainment as a predictor of HIV risk in rural Uganda: Results from a population-based study. International Journal of STD and AIDS 1999;10(7):452-459.

166. Konde-Lule JK, Wawer MJ, Sewankambo NK, Serwadda D, Kelly R, Li C, et al. Adolescents, sexual behaviour and HIV-1 in rural Rakai district, Uganda. AIDS 1997;11(6):791-9.

167. Djamba YK. Premarital sexual experience of married women in Kinshasa, Zaire. Journal of Biosocial Science 1995;27(4):457-66.

168. Grosskurth H, Mosha F, Todd J, Senkoro K, Newell J, Klokke A, et al. A community trial of the impact of improved sexually transmitted disease treatment on the HIV epidemic in rural Tanzania: 2. Baseline survey results. AIDS 1995;9(8):927-934.

169. Quigley M, Munguti K, Grosskurth H, Hodd J, Mosha F, Senkoro K, et al. Sexual behaviour patterns and other risk factors for HIV infection in rural Tanzania: A case-control study. AIDS 1997;11(2):237-248.

170. Kilian AH, Gregson S, Ndyanabangi B, Walusaga K, Kipp W, Sahlmuller G, et al. Reductions in risk behaviour provide the most consistent explanation for declining HIV-1 prevalence in Uganda. AIDS 1999;13(3):391-398.

171. Kirunga CT, Ntozi JP. Socio-economic determinants of HIV serostatus: a study of Rakai District, Uganda. Health Transition Review 1997;7 Suppl:175-188.

172. Mnyika KS, Klepp KI, Kvale G, Ole-King'ori N. Risk factors for HIV-1 infection among women in the Arusha region of Tanzania. Journal of Acquired Immune Deficiency Syndromes 1996;11(5):484-491.

173. Mnyika KS, Klepp KI, Kvale G, Nilssen S, Kissila PE, Ole-King'ori N. Prevalence of HIV-1 infection in urban, semi-urban and rural areas in Arusha region, Tanzania. AIDS 1994;8(10):1477-1481.

174. Lugoe WL, Klepp KI, Skutle A. Sexual debut and predictors of condom use among secondary school students in Arusha, Tanzania. AIDS Care 1996;8(4):443-52.

175. Bassett MT, McFarland WC, Ray S, Mbizvo MT, Machekano R, Van d, et al. Risk factors for HIV infection at enrollment in an urban male factory cohort in Harare, Zimbabwe. Journal of Acquired Immune Deficiency Syndromes 1996;13(3):287-293.

176. Filmer D. The socioeconomic correlates of sexual behavior:A summary of results from an analysis of DHS data. In: Confronting AIDS:Public priorities in a global epidemic. New York: Oxford University Press; 1998.

177. Fylkesnes K, Masonda RM, Kasumba K, Ndhlovu Z, Mluanda F, Kaetano L, et al. The HIV epidemic in Zambia: Socio-demographic prevalence patterns and indications of trends among childbearing women. AIDS 1997;11(3):339-345.

178. Gray RH, Li X, Kigozi G, Serwadda D, Brahmbhatt H, Wabwire-Mangen F, et al. Increased risk of incident HIV during pregnancy in Rakai, Uganda: a prospective study. Lancet 2005;366(9492):1182-8.

179. Bloom SS, Urassa M, Isingo R, Ng'weshemi J, Boerma JT. Community effects on the risk of HIV infection in rural Tanzania. Sexually Transmitted Infections 2002;78:261-266.

180. 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;15:907-916.

181. Meekers D, Ahmed G. Contemporary patterns of adolescent sexuality in urban Botswana. Journal of Biosocial Science 2000;32(4):467-85.

182. Fontanet AL, Woldemichael T, Sahlu T, van Dam GJ, Messele T, Rinke de Wit T, et al. Epidemiology of HIV and Schistosoma mansoni infections among sugar-estate residents in Ethiopia. Annals of Tropical Medicine and Parasitology 2000;94(2):145-55.

183. Kapiga SH, Lugalla JL. Sexual behaviour patterns and condom use in Tanzania: results from the 1996 Demographic and Health Survey. AIDS Care 2002;14(4):455-69.

184. Zuma K, Gouws E, Williams B, Lurie M. Risk factors for HIV infection among women in Carletonville, South Africa: migration, demography and sexually transmitted diseases. International Journal of STD and AIDS 2003;14(12):814-7.

185. Magnani RJ, Karim AM, Weiss LA, Bond KC, Lemba M, Morgan GT. Reproductive health risk and protective factors among youth in Lusaka, Zambia. Journal of Adolescent Health 2002;30(1):76-86.

186. Lurie MN, Williams BG, Zuma K, Mkaya-Mwamburi D, Garnett G, Sturm AW, et al. The impact of migration on HIV-1 transmission in South Africa: a study of migrant and nonmigrant men and their partners. Sexually Transmitted Diseases 2003;30(2):149-56.

187. Koenig MA, Zablotska I, Lutalo T, Nalugoda F, Wagman J, Gray R. Coerced first intercourse and reproductive health among adolescent women in Rakai, Uganda. International Family Planning Perspectives 2004;30(4):156-63.

188. Coffee MP, Garnett GP, Mlilo M, Voeten HA, Chandiwana S, Gregson S. Patterns of movement and risk of HIV infection in rural Zimbabwe. Journal of Infectious Diseases 2005;191 Suppl 1:S159-67.

189. Mbulaiteye SM, Mahe C, Ruberantwari A, Whitworth JA. Generalizability of population-based studies on AIDS: a comparison of newly and continuously surveyed villages in rural southwest Uganda. International Journal of Epidemiology 2002;31(5):961-7.

190. Hallman K. Socioeconomic disadvantage and unsafe sexual behaviours among young women and men in South Africa. New York: Population Council; 2004.

191. Kaufman CE, Clark S, Manzini N, May J. Communities, opportunities, and adolescents' sexual behavior in KwaZulu-Natal, South Africa. Studies in Family Planning 2004;35(4):261-74.

192. Kimuna S, Djamba Y. Wealth and extramarital sex among men in Zambia. International Family Planning Perspectives 2005;31(2):83-9.

193. Taha TE, Dallabetta GA, Hoover DR, Chiphangwi JD, Mtimavalye LA, Liomba GN, et al. Trends of HIV-1 and sexually transmitted diseases among pregnant and postpartum women in urban Malawi. AIDS 1998;12(2):197-203.

194. Evian C, Fox M, MacLeod W, Slotow SJ, Rosen S. Prevalence of HIV in workforces in southern Africa, 2000-2001. South African Medical Journal 2004;94(2):125-30.

195. Booysen Fle R, Summerton J. Poverty, risky sexual behaviour, and vulnerability to HIV infection: evidence from South Africa. Journal of Health, Population and Nutrition 2002;20(4):285-8.

196. Zuma K, Lurie MN, Williams BG, Mkaya-Mwamburi D, Garnett GP, Sturm AW. Risk factors of sexually transmitted infections among migrant and non-migrant sexual partnerships from rural South Africa. Epidemiology and Infection 2005;133(3):421-8.

197. Allen D. South African plans to address HIV/AIDS. In: AIDS in Context, Johannesburg; 2001.

198. Schneider H, Stein J. Implementing AIDS policy in post-apartheid South Africa. Social Science and Medicine 2001;52:723-731.

199. Rogers EM. Diffusion of Innovations. New York: The Free Press; 1962.

200. Mason CJ, Kitsiripornchai S, Markowitz LE, Chanbancherd P, Supapongse T, Jugsudee A, et al. Nationwide surveillance of HIV-1 prevalence and subtype in young Thai men. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology 1998;19(2):165-173.

201. Case A, Deaton A. School inputs and educational outcomes in South Africa. Quarterly Journal of Economics 1999;114(3):1047-1084.

202. Statistics South Africa. Education in South Africa. Selected findings from Census '96. Pretoria: Statistics South Africa; 2001.

203. Rural AIDS and Development Action Research Programme. Social Interventions for HIV/AIDS. Intervention with Microfinance for AIDS and Gender Equity. Evaluation Monograph Number 1. Acornhoek: RADAR; 2002.

204. Rural AIDS and Development Action Research Programme. Social Interventions for HIV/AIDS. Intervention with Microfinance for AIDS and Gender Equity. Intervention Monograph. Number 2. Acornhoek: RADAR; 2002.

205. Hargreaves JR, Pronyk PM, Kim JC, Makhubele MB, Morison LA, Watts C, et al. Protocol 03PRT/24 The Intervention with Microfinance for AIDS and Gender Equity (IMAGE) study: a community randomised trial of a structural intervention to prevent HIV and gender-based violence in South Africa: Lancet Protocol Reviews; 2004.

206. Simanowitz A, Nkuna B, Kasim S. Overcoming the obstacles of identifying the poorest families. Unpublished Report; 2000.

207. Simanowitz A, Nkuna B. Participatory Wealth Ranking Operational Manual. Tzaneen: Small Enterprise Foundation; 1998.

208. Hargreaves JR, Morison LA, Gear J, Porter JDH, Makhubele MB, Kim JC, et al. "Hearing the voices of the poor": Assigning poverty lines on the basis of local perceptions of poverty; a quantitative analysis of qualitative data from participatory wealth ranking in rural South Africa. World Development 2005;In Press.

209. Gregson S, Nyamukapa CA, Garnett GP, Mason PR, Zhuwau T, Carael M, et al. Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. Lancet 2002;359(9321):1896-903.

210. Muller R, Buttner P. A critical discussion of intraclass correlation coefficients. Statistics in Medicine 1994;13:2465-2476.

211. Lee J, Koh D, Ong CN. Statistical evaluation of agreement between two methods for measuring a quantitative variable. Comput Biol Med 1989;19(1):61-70.

212. Kirkwood B, Sterne J. Analysis of clustered data. In: Essential Medical Statistics. 2nd ed. Oxford: Blackwell Publishing; 2003. p. 355-370.

213. Victora CG, Huttly SR, Fuchs SR, Olinto MTA. The Role of Conceptual Frameworks in Epidemiological Analysis. International Journal of Epidemiology 1997;26(1):224-227.

214. Gregson S, Mushati P, White PJ, Mlilo M, Mundandi C, Nyamukapa C. Informal confidential voting interview methods and temporal changes in reported sexual risk behaviour for HIV transmission in sub-Saharan Africa. Sexually Transmitted Infections 2004;80 Suppl 2:ii36-42.

215. Gregson S, Zhuwau T, Ndlovu J, Nyamukapa CA. Methods to reduce social desirability bias in sex surveys in low-development settings: experience in Zimbabwe. Sexually Transmitted Diseases 2002;29(10):568-75.

216. Buve A, Lagarde E, Carael M, Rutenberg N, Ferry B, Glynn JR, et al. Interpreting sexual behavior data: validity issues in the multicentre study on factors determining the differential spread of HIV in four African cities. AIDS 2002;15 (suppl 4):S117-S126.

217. UNGASS. Declaration of Commitment on HIV/AIDS. In: United Nations General Assembly Special Session on HIV/AIDS; 2001; New York: United Nations; 2001.

218. Kaufman C, de Wet T, Stadler J. Adolescent pregnancy and parenthood in South Africa. New York: Population Council; 2000.

219. Garenne M, Tollman S, Kahn K. Premarital fertility in rural South Africa: a challenge to existing population policy. Stud Fam Plann 2000;31(1):47-54.

220. Haddad L, Hoddinott J, Alderman H. Intrahousehold resource allocation in developing countries. Baltimore: Johns Hopkins University Press; 1997.

221. Bergeron G, Morris SS, Medina Banegas JM. How reliable are group informant ratings? A test of food security ratings in Honduras. World Development 1998;26(10):1893-1902.

222. Posel DR. Intra-family transfers and income-pooling: A study of remittances in Kwazulu-Natal. South African Journal of Economics 2001;69(3):501-528.

223. Catania JA, Binson D, Canchola J, Pollack LM, Hauck W, Coates TJ. Effects of interviewer gender, interviewer choice, and item wording on responses to questions concerning sexual behavior. Public Opinion Quarterly 1996:345-375.

224. Catania JA, Gibson DR, Chitwood DD, Coates TJ. Methodological Problems in AIDS behavioural research:influences on measurement error and participation bias in studies of sexual behavior. Psychological Bulletin 1990;108(3):339-362.

225. Pequegnat W, Fishbein M, Celentano D, Ehrhardt A, Garnett GP, Holtgrave D, et al. NIMH/APPC Workgroup on behavioral and biological outcomes in HIV/STD prevention studies. A position statement. Sexually Transmitted Diseases 2000;27(1):127-132.

226. Jewkes R, Watts C, Abrahams N, Penn-Kekana L, Garcia-Moreno C. Ethical and methodological issues in conducting research on gender-based violence in Southern Africa. Reprod Health Matters 2000;8(15):93-103.

227. Gallo D, George R, Fitchen J, Goldstein A, Hindahl MS. Evaluation of a System using Oral Mucosal Transudate for HIV-1 Antibody Screening and confirmatory testing. Journal of the American Medical Association 1997;277:254-258.

228. Holm-Hansen C. Detection of antibodies to HIV in homologous sets of plasma, urine and oral mucosal transudate samples using rapid assays in Tanzania. Clinical and Diagnostic Virology 1993;1:207-214.

229. Hargreaves JR, Stevens G, Makhuraj A, Morison LA, Porter JDH, Watts C, et al. Preliminary estimates of the acceptability of oral fluid collection, and the sensitivity, specificity and reproducibility of ELISA analysis of oral fluid samples for detection of antibodies to HIV from a field study in South Africa. In: South African AIDS Conference. Durban; 2003.

230. Gregson S, Zhuwau T, Anderson RM, Chimbadzwa T, Chiwandiwa SK. Age and religion selection biases in HIV-1 prevalence data from antenatal clinics in Manicaland, Zimbabwe. Central African Journal of Medicine 1995;41(11):339-346.

231. Serwadda D, Gray RH, Wawer MJ, Stallings RY, Sewankambo NK, Konde-Lule JK, et al. The social dynamics of HIV transmission as reflected through discordant couples in rural Uganda. AIDS 1995;9:745-750.

232. Batty GD, Deary IJ. Education and mortality: the role of intelligence. Lancet 2005;365(9473):1765-6.

233. Crewe M. Socioeconomic determinants of HIV and AIDS in South Africa. In: National STD/HIV/AIDS Review. Johannesburg: Medical Research Council, South Africa; 1997.

234. Evian C. AIDS and the socio-economic determinants of the epidemic on Southern Africa - A cycle of poverty. Journal of Tropical Pediatrics 1994;40:61-62.

235. Rose G. Sick individuals and sick populations. International Journal of Epidemiology 2001;30(3):427-32; discussion 433-4.

236. Meekers D, Calves AE. 'Main' girlfriends, girlfriends, marriage, and money: the social context of HIV risk behaviour in sub-Saharan Africa. Health Transition Review 1997;7 Suppl:361-75.

237. Peters C, Makhubele MB, Phetla G, Hargreaves JR, Kim JC, Pronyk PM. You'll die and leave the polonies in the fridge. In: Sex and Secrecy. Johannesburg; 2003.

238. Abdelrahman AI. Education and assortative marriage in northern and urban Sudan: 1945-79. Journal of Biosocial Science 1994;26(3):341-8.

239. Doherty IA, Padian NS, Marlow C, Aral SO. Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. Journal of Infectious Diseases 2005;191 Suppl 1:S42-54.

240. Kelly RJ, Gray RH, Sewankambo NK, Serwadda D, Wabwire-Mangen F, Lutalo T, et al. Age differences in sexual partners and risk of HIV-1 infection in rural Uganda. Journal of Acquired Immune Deficiency Syndromes 2003;32(4):446-51.

241. Glynn JR, Carael M, Auvert B, Kahindo M, Chege J, Musonda R, et al. Why do young women have a much higher prevalence of HIV than young men? A study in Kisumu, Kenya and Ndola, Zambia. Aids 2001;15 Suppl 4:S51-60.

242. Luke N. Confronting the 'sugar daddy' stereotype: age and economic asymmetries and risky sexual behavior in urban Kenya. International Family Planning Perspectives 2005;31(1):6-14.

243. HIV/AIDS&STD: Strategic Plan for South Africa, 2000-2005: Department of Health, South Africa; 2000.

244. Pronyk PM, Kim JC, Makhubele MB, Hargreaves JR, Mohlala R, Hausler HP. Introduction of voluntary counselling and rapid testing for HIV in rural South Africa: from theory to practice. AIDS Care 2002;14(6):859-65.

245. Magnani R, Macintyre K, Karim AM, Brown L, Hutchinson P, Kaufman C, et al. The impact of life skills education on adolescent sexual risk behaviors in KwaZulu-Natal, South Africa. Journal of Adolescent Health 2005;36(4):289-304.

246. Parker W. Re-appraising youth prevention: the case of Lovelife in South Africa. In: 1st South African AIDS Conference. Durban, South Africa; 2003.

247. Blakenship KM, Bray SJ, Merson MH. Structural interventions in public health. AIDS 2000;14 (Suppl 1):S11-S21.

248. Fenton L. Preventing HIV/AIDS through poverty reduction: the only sustainable solution? Lancet 2004;364(9440):1186-7.

249. Lemon A. Redressing school inequalities in the Eastern Cape, South Africa. Journal of Southern African Studies 2004;30(2):269-290.

250. Department of Education. Review of the financing, resourcing and costs of education in public schools. Pretoria: Department of Education, South Africa; 2003.

251. Morris SS, Flores R, Olinto P, Medina JM. Monetary incentives in primary health care and effects on use and coverage of preventive health care interventions in rural Honduras: cluster randomised trial. Lancet 2004;364(9450):2030-7.

252. Rivera JA, Sotres-Alvarez D, Habicht JP, Shamah T, Villalpando S. Impact of the Mexican program for education, health, and nutrition (Progresa) on rates of growth and anemia in infants and young children: a randomized effectiveness study. Journal of the American Medical Association 2004;291(21):2563-70.

253. Johnson S, Rogaly B. Microfinance and poverty reduction. London: Oxfam 1997.

254. Sebstad J, Chen G. Overview of studies on the impact of microenterprise credit. Washington: USAID; 1996.

255. Sebstad J, Neill C, Barnes C, Chen G. Assessing the impacts of microenterprise interventions. A framework for analysis. Arlington: USAID; 1995.

256. Todd H. Women at the Centre. Westview Press: Oxford; 2002.