

---

# AGE DIFFERENCES IN SEXUAL RELATIONSHIPS AND HIV TRANSMISSION: STATISTICAL ANALYSES OF BIO- BEHAVIOURAL SURVEY DATA FROM SOUTHERN AFRICA

---

Roxanne Beauclair

*Doctoral thesis submitted to the Faculty of Medicine and Health Sciences, Ghent  
University*

October 2017

**PhD Supervisors:**

Prof. dr. Wim Delva  
International Centre for Reproductive Health (ICRH)  
Dept. of Uro-gynecologie  
Ghent University, Ghent, Belgium

Prof. dr. Niel Hens  
Center for Statistics  
Hasselt University, Diepenbeek, Belgium

Prof. dr. Marleen Temmerman  
International Centre for Reproductive Health (ICRH)  
Dept. of Uro-gynecologie  
Ghent University, Ghent, Belgium

Age Differences in Sexual Relationships and HIV Transmission: Statistical Analyses of  
Bio-Behavioural Survey Data from Southern Africa

Roxanne Beauclair

Doctoral thesis submitted to the Faculty of Medicine and Health Sciences, Ghent  
University

October 2017

ISBN: 9789078128519

EAN: 9789078128519

International Centre for Reproductive Health (ICRH)  
Ghent University  
De Pintelaan 185P3  
B-9000 Ghent (Belgium)  
[www.icrh.org](http://www.icrh.org)

### **PhD Supervisors:**

Prof. dr. Wim Delva  
International Centre for Reproductive Health (ICRH)  
Dept. of Uro-gynecologie  
Faculty of Medicine and Health Sciences  
Ghent University, Ghent, Belgium

Prof. dr. Niel Hens  
Center for Statistics  
Hasselt University, Diepenbeek, Belgium

Prof. dr. Marleen Temmerman  
International Centre for Reproductive Health (ICRH)  
Dept. of Uro-gynecologie  
Faculty of Medicine and Health Sciences  
Ghent University, Ghent, Belgium

### **Chairperson of the jury:**

Prof. dr. Olivier DeGomme  
International Centre for Reproductive Health (ICRH)  
Ghent University

### **Members of the jury:**

Prof. dr. Kristien Michiels  
International Centre for Reproductive Health (ICRH)  
Ghent University

Prof. dr. Els Clays  
Department of Public Health  
Ghent University

Prof. dr. Stijn Vansteelandt  
Department of Applied Mathematics, Computer Science and Statistics  
Ghent University

Prof. dr. Ziv Shkedy  
Center for Statistics  
Hasselt University

Dr. Chris Kenyon  
The Institute of Tropical Medicine  
University of Antwerp

Dr. Brendan Maughan-Brown  
Southern Africa Labour and Development Research Unit (SALDRU)  
University of Cape Town

# Contents

- List of abbreviations ..... 1
- Preface ..... 2
- CHAPTER 1: INTRODUCTION ..... 4**
  - 1.1 Context and overview ..... 5
  - 1.2 Commonly used terminology and definitions ..... 6
    - 1.2.1 Age differences ..... 6
    - 1.2.2 Measures of occurrence and association ..... 8
  - 1.3 Age differences in the context of HIV research ..... 9
    - 1.3.1 Age and gender-stratified HIV prevalence ..... 10
    - 1.3.2 Sexual behaviour correlates of age-disparate relationships ..... 11
    - 1.3.3 Biological susceptibility of young women to HIV infection ..... 12
    - 1.3.4 Reasons for pursuing age-asymmetric relationships ..... 13
  - 1.4 Association between age differences and HIV infection ..... 14
    - 1.4.1 Dynamical models ..... 14
    - 1.4.2 Ecological associations of age-asymmetry and HIV prevalence ..... 15
    - 1.4.3 Individual-level observational studies of age-asymmetry and HIV infection ..... 16
  - 1.5 Motivation ..... 18
  - 1.6 Objectives ..... 19
  - 1.7 Thesis structure ..... 19
- CHAPTER 2: METHODS ..... 20**
  - 2.1 Study design and overview ..... 21
  - 2.2 An abridged description of the datasets ..... 21
    - 2.2.1 Cape Area Panel Study (CAPS) ..... 21
      - 2.2.1.1 Study synopsis ..... 21
      - 2.2.1.2 Study setting ..... 21
    - 2.2.2 Likoma Network Study (LNS) ..... 22
      - 2.2.2.1 Study synopsis ..... 22
      - 2.2.2.2 Study setting ..... 23
    - 2.2.3 Schooling, Income and Health Risk (SIHR) Study ..... 24
      - 2.2.3.1 Study synopsis ..... 24
      - 2.2.3.2 Study setting ..... 26
    - 2.2.4 Cape Town Sexual Behaviour Study (CTSBS) ..... 26
      - 2.2.4.1 Synopsis of the sexual behaviour survey ..... 26
      - 2.2.4.2 Synopsis of the qualitative Interviews ..... 28
      - 2.2.4.3 Study setting ..... 29
  - 2.3 Frequently used analytical methods ..... 29
    - 2.3.1 Regression modelling ..... 29
    - 2.3.2 Multilevel models ..... 31

2.3.3	Non-linear covariates .....	32
2.3.4	Multivariate Imputation by Chained Equations .....	33
2.3.5	Thematic analysis .....	35
2.4	Ethics approval and considerations .....	36
2.5	Dissemination of results .....	37
<b>CHAPTER 3: AGE-DISPARATE RELATIONSHIPS AND IMPLICATIONS FOR STI TRANSMISSION AMONG YOUNG ADULTS IN CAPE TOWN, SOUTH AFRICA .....</b>		
3.1	Aims of analysis .....	40
3.2	Manuscript .....	41
<b>CHAPTER 4: IS YOUNGER REALLY SAFER? A QUALITATIVE STUDY OF PERCEIVED RISKS AND BENEFITS OF AGE-DISPARATE RELATIONSHIPS AMONG WOMEN IN CAPE TOWN, SOUTH AFRICA .....</b>		
4.1	Aims of analysis .....	52
4.2	Manuscript .....	53
<b>CHAPTER 5: AGE DIFFERENCES BETWEEN SEXUAL PARTNERS, BEHAVIOURAL AND DEMOGRAPHIC CORRELATES, AND HIV INFECTION ON LIKOMA ISLAND, MALAWI .....</b>		
5.1	Aims of analysis .....	62
5.2	Manuscript .....	63
<b>CHAPTER 6: PARTNER AGE DIFFERENCES AND ASSOCIATED SEXUAL RISK BEHAVIOURS AMONG ADOLESCENT GIRLS AND YOUNG WOMEN IN A CASH TRANSFER PROGRAMME FOR SCHOOLING IN MALAWI .....</b>		
6.1	Aims of analysis .....	79
6.2	Manuscript .....	80
<b>CHAPTER 7: THE ROLE OF AGE-MIXING PATTERNS IN HIV TRANSMISSION DYNAMICS: NOVEL HYPOTHESES FROM A FIELD STUDY IN CAPE TOWN SOUTH AFRICA .....</b>		
7.1	Aims of analysis .....	103
7.2	Manuscript .....	104
<b>CHAPTER 8: DISCUSSION .....</b>		
8.1	Chapter precis .....	135
8.2	Epidemiological implications of key findings .....	135
8.2.1	Risky sexual behaviours correlated with large age differences .....	135
8.2.2	Age-mixing may create conditions necessary to sustain HIV in the population .....	135
8.2.3	Relationship between age differences and prevalent HIV infection in individuals .....	137
8.2.4	Psycho-social determinants of age differences in relationships .....	138

8.2.5	The role of structural interventions in reducing age differences .....	138
8.3	Limitations.....	139
8.4	Research recommendations .....	141
8.4.1	Improvement in the design of sexual behaviour studies.....	141
8.4.2	Analytical recommendations .....	142
<b>CHAPTER 9: CONCLUSIONS .....</b>		<b>144</b>
9.1	Recommendations for public health interventions .....	145
9.2	Conclusions .....	147
<b>ABSTRACT .....</b>		<b>148</b>
<b>SAMENVATTING.....</b>		<b>150</b>
<b>ACKNOWLEDGEMENTS .....</b>		<b>153</b>
<b>LIST OF PUBLICATIONS.....</b>		<b>156</b>
<b>APPENDIX A .....</b>		<b>157</b>
<b>APPENDIX B .....</b>		<b>164</b>
<b>APPENDIX C.....</b>		<b>171</b>
<b>APPENDIX D .....</b>		<b>179</b>
<b>APPENDIX E.....</b>		<b>188</b>
<b>REFERENCES.....</b>		<b>190</b>

## *Figures*

Figure 1.	Factors associated with HIV transmission and large age differences .....	9
Figure 2.	HIV Prevalence in South Africa by age group and gender .....	10
Figure 3.	HIV Prevalence in Malawi by age group and gender .....	11
Figure 4.	Likoma Network Study Participant Flow Chart .....	23
Figure 5.	Schooling, Income and Health Risk Study Participation Flow Chart .....	25
Figure 6.	Demonstration of ACASI on touchscreen computer .....	28

## *Tables*

Table 1.	Definitions for age difference terms.....	7
Table 2.	Summary of regression models used in the results chapters .....	31

## List of abbreviations

---

ACASI	Audio Computer-Assisted Self-Interviewing
ADR	Age-disparate Relationship
AIDS	Acquired Immunodeficiency Syndrome
ART	Anti-retroviral Therapy
CAPS	Cape Area Panel Study
CC	Complete Cases
CI	Confidence Interval
CPHM	Cox Proportional Hazard Model
CTSBS	Cape Town Sexual Behaviour Study
DHS	Demographic and Health Surveys
DMPA	Depot Medroxyprogesterone Acetate
EBWR	Expected Bridge Width Ratio
FTFI	Face-to-face Interviewing
GAM	Generalized Additive Models
GLM	Generalized Linear Models
GLMM	Generalized Linear Mixed Models
HIV	Human Immunodeficiency Virus
HR	Hazard Ratio
IGR	Inter-generational Relationship
IHC	Injectable Hormonal Contraceptive
LNS	Likoma Network Study
MAR	Missing at Random
MCAR	Missing Completely At Random
MI	Multiple Imputation
MICE	Multivariate Imputation by Chained Equations
MNAR	Missing Not At Random
OR	Odds Ratio
POR	Proportional Odds Ratio
RF	Random Forrest
RR	Risk Ratio
SE	Standard Error
SIHR	Schooling, Income and Health Risk
STI	Sexually Transmitted Infection
TB	Tuberculosis

# Preface

---

The dissertation presented here addresses an important and urgent matter: the increasing amount of new HIV infections among young women in Southern Africa, and whether age disparities in the sexual relationships are driving this feature of the epidemic. Before I delve into the context, motivation, and hypotheses of my research, I would like to take a slight detour and briefly describe my journey for how I arrived at this point and lessons learned along the way. My story has less to do with the topic of HIV infection, and more to do with the (re)discovery of my love for math, algorithms, simplicity, design and context.

Nearly 10 years ago I found myself in Cape Town, South Africa at the commencement of a graduate programme in social anthropology. As a budding, young American anthropologist, I found myself incredibly interested in anything having to do with “other cultures”. I was originally driven by disparate questions (no pun intended) about epistemological frameworks, kinship structures, social constructions of identity, as well as other “exotic” beliefs, like witchcraft. These academic interests gradually gave way to my more focused fascination with systems of medical knowledge and care, as well as a critical engagement in development programmes to eliminate poverty and provide services to marginalized communities. As a student of anthropology, I quickly learned that: the world is messy; nothing is universal; and context is everything. Yet, I found myself immobilized by these realizations and I was left wondering if there was anything I could personally do to improve the quality of life for people in need of aid and resources.

I decided to take the best parts of my anthropological education – to be critical of things deemed as “natural”; to question systems of power and exclusion; and to understand the impact of interventions and new technologies in non-Western settings – and thus, I began a new Master’s programme in Public health (MPH). For the first time in nearly a decade I found myself enrolled in quantitative modules. At first, I did not really understand what I was getting myself into, but soon I came to enjoy my biostatistics courses the most. Different muscles in my brain were exercised and I realized how much I had previously been missing the rules, order, and explicit assumptions that the field of statistics brought into to my life. I no longer felt crippled by my fear of becoming a misguided “white saviour” in the field of global health. I felt empowered because I was learning an in-demand skillset that anyone in any discipline or global setting could appreciate.

Nearing the end of my MPH, I received the opportunity to work with my current promoter, Prof. Wim Delva, at the South African Centre for Epidemiological Modelling & Analysis (SACEMA). Though I was only offered a job as the study coordinator of a sexual behaviour survey, I jumped at the opportunity because I was secretly hoping it would lead to more hands-on experience with analysis of epidemiological data. It was during that role that I was first introduced to my other two co-promoters, Prof. Niel Hens and Prof. Marleen Temmerman, though at the time they were professional collaborators. In the end, I probably got more than I bargained for when Prof. Delva hired me, because very soon after, he started encouraging me to write papers (relentlessly), and eventually, conduct my own analyses. As I was gaining this practical

experience, I developed a new appreciation for elegant, uncomplicated, and coherent statistical programming.

After three years of working at SACEMA, I jumped at the chance to start my PhD at Ghent University when we successfully applied for and received funding for the work we had already begun. This PhD work has taught me several lessons that I hope to take with me into my future positions. First, though there are statistical methods that can mitigate confounding and bias, there are no statistical methods that can correct for all of the flaws present in a poorly designed study. Second, novel and advanced statistical methods always look promising in the theoretical paper, but when they are confronted with the ugly mess of real-world data they often fall apart. Third, the rule of parsimony does not only apply to statistical models themselves. Sometimes during the workflow of a statistical analysis, one has to consider settling for a simple, approximate, and economical solution to an analytical problem.

Finally, in many ways I feel as though my journey has come full circle. I have realized that there is no such thing as a statistical analysis that does not violate some of the assumptions made by the chosen methods and that uses perfectly collected data free from bias. Like the realization I came to at the completion of my anthropological adventure, each analysis is situated within layers of contexts that can be chaotic at times, but still need consideration. An analyst must consider the study design context, in order to choose appropriate methods of analysis; be cognizant of the funding context for the sake of adhering to protocol and reporting deadlines; and be aware of the political, social and geographical contexts to give meaning to and appropriately interpret the results.

These lessons have been difficult, painstakingly slow to learn, and at times extremely frustrating, but alas, there they are. Though these specific lessons will not be discussed again in my *magnum opus*, I hope that as you continue to read they are apparent in my motivations and study implications. Thank you for sharing in this journey with me.

Roxanne Beauclair

Sept 11, 2017

## CHAPTER 1: INTRODUCTION

---



Photo credit: IMs Bildarkiv / Photo Archive via Visual Hunt / CC BY-NC-ND

## 1.1 CONTEXT AND OVERVIEW

---

In 2015 there were 36.7 million adults and children living with HIV in the world, with approximately 19 million of those living in Southern and Eastern Africa (1). Specifically, in South Africa and Malawi the HIV epidemic is characterized by heterosexual transmission and the HIV prevalences among adults aged 15-49 years old in 2015 were 19.2% and 9.1%, respectively (2).

The evolution of the HIV epidemic in South Africa has been widely documented since the first case was reported in 1982. Until 1990, the prevalence of HIV was relatively low and was observed mostly in mineworkers, voluntary blood donors, and men who have sex with men (3). In 1990, for the first time, HIV was reported in heterosexual women who attended public antenatal care clinics and the prevalence was estimated to be approximately 1% (3). Since then, the national HIV prevalence in pregnant women has risen exponentially: from 7.6% in 1994, to 20.5% in 2000, and peaking at 30.2% in 2005 (4). In 2006, the HIV incidence rate was estimated to be in excess of 5% per year (4). Though incidence was on the rise in the mid-2000s, the increasing mortality rate meant that there were small reductions in the overall prevalence. The HIV epidemic in Malawi has not been described as well as in South Africa, although it has taken a similar trajectory. In Malawi, the prevalence rapidly increased from 1990 to 2002, where thereafter it stabilized (5). By 2000, the incidence of HIV was already starting to decrease (5).

Great strides in improving access to anti-retroviral therapy (ART) in South Africa and Malawi have been made since the mid-2000s, and as of 2015, 61% of those infected in Malawi, and 48% infected in South Africa were receiving treatment (2). ART has the ability to virally suppress HIV positive individuals and thus prevent secondary transmissions to sexual partners who are HIV negative (6,7). Despite the successes with ART coverage in these countries, there are still a substantial number of new infections each year. In South Africa the estimated HIV incidence among 15-49 year olds was 1.44 infections per 100 person-years and 0.38 infections per 100 person-years in Malawi (2).

Though adolescent and young women aged 15-24 years comprise only 11% of the global adult population, they account for 20% of new HIV infections (1). In sub-Saharan Africa the disparity is more striking, with women in this age group constituting 25% of new infections, while young men and adolescent boys make up only 12% of new infections (1). Thus, young women and adolescent girls have been identified as a key population for targeting HIV prevention interventions (8,9). Large age-differences in sexual relationships have been proposed as an explanation for the gender imbalance in HIV incidence among young adults, as well as for the magnitude and persistence of the epidemic in Southern Africa (10-13).

Through four secondary statistical analyses of epidemiological data, as well as a qualitative study of relationship dynamics in Cape Town, I intend to demonstrate how age differences in heterosexual relationships have the potential to both, increase individual-level risk of HIV transmission, and help maintain the epidemic within populations of South Africa and Malawi. In this chapter, I will present the current state of research on this topic. First, I will provide definitions for terminology commonly used throughout the dissertation, followed by a more in-depth background on why age differences in relationships are suspected to play an important role in HIV transmission. Then I will assess the evidence for the claim that large age differences in relationships

are correlated with HIV infection by summarizing previous epidemiological studies on the subject and critiquing their methods. Next, I will provide a motivation for my own research by highlighting the still-unknown questions in this field. Finally, I will end the chapter by outlining the over-arching aims of my research and the specific manuscripts that have helped to achieve those objectives.

## 1.2 COMMONLY USED TERMINOLOGY AND DEFINITIONS

---

### 1.2.1 Age differences

The most fundamental term underlying nearly all of the literature and research that will be discussed in this dissertation is the notion of an *age difference*. Age differences are defined as the age of a person minus the age of their sexual partner, and they are expressed in years. Though it may seem obvious, it is important to note that the ages of participants and their partners are rounded down to the nearest year. In other words, a person's age is always a whole number and not expressed as a fraction. In most cases, the female partner age is subtracted from the male partner's age, regardless of whether it was the man or the woman who was the actual participant in the study. When a negative age difference is presented (e.g.  $-5$  years), it means that the female partner was older than the male partner by that many years. Age differences are relationship-level characteristics, and thus a single participant in a study may have as many age differences connected to them as the number of relationships they report. Age gaps are synonymous with age differences, and may be used interchangeably throughout the dissertation.

There are several other terms related to age differences, which are presented in different studies. Perhaps the most general term is *age-asymmetry*. When it is used it generally refers to any relationship in which the age difference is not zero, i.e. the male and female partner are not the same age. An *age-disparate relationship* (ADR) is the dichotomized version of an age difference, and it is usually defined as a relationship in which the male partner is 5 or more years older than the female partner (14). A variation of this is the *Intergenerational Relationship* (IGR), which is a relationship in which the age difference is 10 or more years (14).

When researchers want to refer to population-level patterns for how people choose partners with regards to age they use the term, *age-mixing pattern*. For example, imagine a population where the average age difference is 7 years old, and as individuals grow older they tend to have wider variation in the ages of partners that they choose. These are characteristics of the population, rather than a relationship or individual. When one describes a population in which most people tend to choose age-similar partners, they may use the term *age-assortative* (15). Likewise, an *age-disassortative* population is one in which people, on average, choose age-asymmetric relationships.

For some of the new research presented in this dissertation, my collaborators and I coined the term, *bridge width*. We hypothesized that in order for HIV to be maintained in a population over a long period of time, HIV would have to be transmitted back and forth between birth cohorts. Commonly used measures such as, ADR or IGR, do not fully capture this idea and are typically used only as an indicator of an individual's risk of acquiring HIV. Thus, we conceptualized *bridge width* as an individual-level indicator of a person's ability to transmit HIV to different age groups. This term is defined as the

number of years difference between the maximum and minimum age difference (i.e. the range of age differences) from partners that were reported in a specified period of time. It can also be technically defined as the difference between the maximum partner age and the minimum partner age. Individuals who have a large range of age differences (or partner ages), would be better situated to act as a “bridge” for HIV to enter other age groups.

As an example, if a woman who was 20 years old at the time of the survey reported two partners: one who was 50 years old and another partner 18 years old at the time of the survey, her *bridge width* would be 32. If a 25-year-old man reported four partners aged 21, 22, 24, and 25, his *bridge width* would be 4. In this example, the man has more relationships, but his potential to transmit HIV between birth cohorts would be less. It is crucial to note that a *bridge width* may not be a good determinant of an individual’s risk of acquiring HIV, but rather its importance relates to the potential for maintaining a high HIV prevalence in a population.

A quick reference to all definitions for age difference and related terms can be found in Table 1.

<i>Term</i>	<i>Definition</i>	<i>Unit of observation to which the term applies</i>
<i>Age difference, age gap</i>	Equals the age of the person minus the age of their sexual partner	Relationship
<i>Age-asymmetry</i>	Describes a relationship in which the age difference is not zero	Relationship
<i>Age-similar</i>	Describes a relationship in which the age of both partners is the same, or within a few years of each other	Relationship
<i>Age disparate relationship</i>	Typically refers to a relationship where the male partner is 5 or more years older than the female partner	Relationship
<i>Intergenerational/Cross-generational relationship</i>	A relationship where the male partner is 10 or more years older than the female partner	Relationship
<i>Bridge width</i>	The number of years difference between the maximum and minimum age difference from partners that were reported in a specified period of time.	Participant
<i>Age-mixing pattern</i>	Population-level trends for how individuals choose partners, with regards to age	Population
<i>Age-assortative</i>	Refers to people choosing age-similar partners	Population
<i>Age-disassortative</i>	Refers to people choosing age-asymmetric partners	Population
<i>Sugar daddy</i>	Colloquial term used to describe older male partners of young women. The connotation is that relationships are typically transactional.	Male partner
<i>Blesser</i>	Colloquial term used to describe older male partners in South Africa. The connotation is that relationships are typically transactional.	Male partner

**Table 1. Definitions for age difference terms**

## 1.2.2 Measures of occurrence and association

In the research that follows, the term *prevalence* refers to the proportion of people in a population who have a particular outcome of interest. It is expressed as a fraction and calculated by dividing the number of people who were found to have the outcome of interest by the number of people in the study population. For the purposes of this work, prevalences (point prevalences, to be specific) were measured and reported when the study design was cross-sectional, or the fraction of people who had the outcome at a specific point in time was of interest. This contrasts with *incidence*, which is a measure of how many new occurrences of the outcome happened in a specified amount of exposure time. It is a rate consisting of the number of new events in the numerator and the time at risk (total time observed for all participants at risk in the study) in the denominator. This can only be measured if the study design is longitudinal and participants are followed-up over time. Similarly, a *risk* is defined as the proportion of participants in the study population – initially free of outcome of interest – who experience the outcome of interest. Less commonly used as a measure of occurrence is an *odds*. This is defined as the ratio of the number of people who experienced the outcome to the number of people in a study who did not experience the outcome. These measures of occurrence are used to describe the frequency of HIV and ADR in much of the literature.

For the analyses in this dissertation, I was primarily concerned with obtaining and reporting measures of association, rather than the measures of occurrence defined above. The statistical models that were employed produced *odds ratios*, *proportional odds ratios*, *risk ratios*, or *hazard ratios*. All of these ratios quantify the relationship between an exposure and a health outcome. Building on the definition of odds above, an odds ratio (OR) is the ratio of odds of having an outcome in those exposed to the odds of having the outcome in those not exposed. If the outcome of interest has more than two categories and it is ordered, it is preferable to report a proportional odds ratio (POR). This measure describes the odds of experiencing all of the higher-ordered categories of the outcome versus the lower-ordered categories for a particular exposure. Concretely, if for example, one wanted to determine if ADRs (exposure) were associated with condom use – an ordinal outcome (1. Never, 2. Inconsistently, or 3. Always) – a POR could be calculated. It would indicate whether the odds of using condoms “Inconsistently/Always” versus “Never” were different for those in ADRs versus age-similar relationships. Likewise, the POR would tell you the odds of using condoms “Always” versus “Inconsistently/Never”.

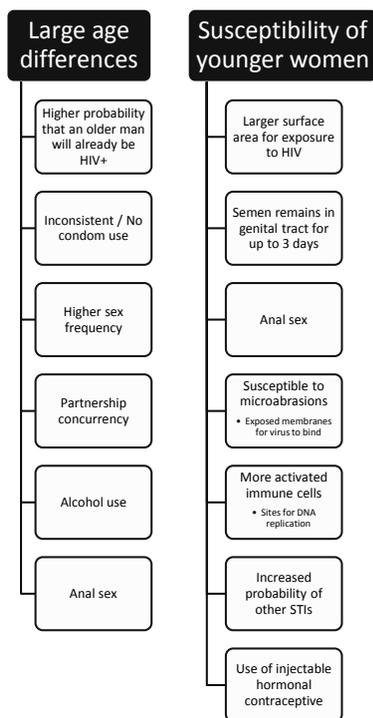
For many people, ORs are not intuitive and thus risk ratios (RRs) are the preferred measures of association. A RR compares the risk of the outcome among one group with the risk of experiencing the outcome in another group. For example, if 20% of women who have an ADR are HIV positive, and only 10% of women who do not have an ADR are HIV positive, then the RR would be 2.00. In other words, women in an ADR have 2 times the risk of being HIV infected. While RRs are preferred, they are often more difficult to obtain, and thus when the outcome of interest is rare in a population an OR can be substituted and be interpreted as an RR because they will be approximately equal (16,17). Finally, a hazard ratio (HR) is similar to a RR, but differs in that it represents the instantaneous risk of experiencing an event (i.e. in the mathematical limit of the time window of being at risk of the event going to zero) in one exposure

group compared to another. The measure of association reported in a study largely depends on the types of variables and nature of the data under consideration.

The measures of occurrence and association described here are fairly common in epidemiological studies and more information about them can be found in the seminal works of Rothman, Greenland, and Lash (18). More details about the statistical models that were used in this dissertation and which measures of association they produced are featured in Section 2.3.1.

### 1.3 AGE DIFFERENCES IN THE CONTEXT OF HIV RESEARCH

It was previously noted that young women account for a disproportionate share of new HIV infections in the world, and in sub-Saharan Africa more specifically. In this section, I will elucidate why researchers suspect that age differences in relationships may be partially responsible for this disparity. To do this, I will describe the different risk factors which may put young women at increased risk of acquiring HIV when they are in relationships with older men, as well as reasons why young women engage in relationships with older men. A schematic overview of these risk factors can be viewed in Figure 1.



**Figure 1. Factors associated with HIV transmission that are also correlated with large age differences and the increased susceptibility of younger women.**

### 1.3.1 Age and gender-stratified HIV prevalence

A 2012 nationally-representative survey of HIV prevalence and incidence in South Africa demonstrated that 5.6% of 15-19-year-old women and 17.4% of 20-24-year-old women were HIV positive (19). This compared to 0.7% HIV prevalence in 15-19-year-old men and 5.1% HIV prevalence in 20-24-year-old men (19). In other words, adolescent girls and young women in these age groups were roughly 8 and 3 times more likely to be HIV positive than their male peers, respectively. The same study indicated that HIV prevalence peaked for men in the 35-39-year age group at 28.8% (see Figure 2) (19).

In the Demographic and Health Survey (DHS) report from Malawi in 2010, similar trends were observed (see Figure 3). In the 15-19-year-old age group the HIV prevalence was 4% and 1% in women and men, respectively (20). In both men and women, the prevalence grew with age and peaked at 24% for women in the 35-39 year age group and 21% for men in the 40-44 year age group (20). More recent age and gender-stratified estimates of prevalence in Malawi were not available.

In South Africa roughly one-third of young women aged 15-19 years old reported an ADR in 2012 (19). In Malawi, only 1% of women in this age group reported an IGR during the DHS survey cited above (20). Regrettably, nationally-representative estimates of ADR in Malawi for this age group are unavailable, though in one smaller, isolated study population the estimated fraction of relationships that were ADR was 15.6% in 18-49 year olds (21). These ADR prevalence estimates, taken together with the age and gender HIV prevalence trends imply that if young women in both countries were selecting older men to have sexual relationships with, they would have had a greater likelihood of choosing a partner who was HIV positive. Whereas, age-similar young, male partners would have been sexually active for a shorter period of time, and thus have a smaller probability of being HIV positive (19,20,22,23).

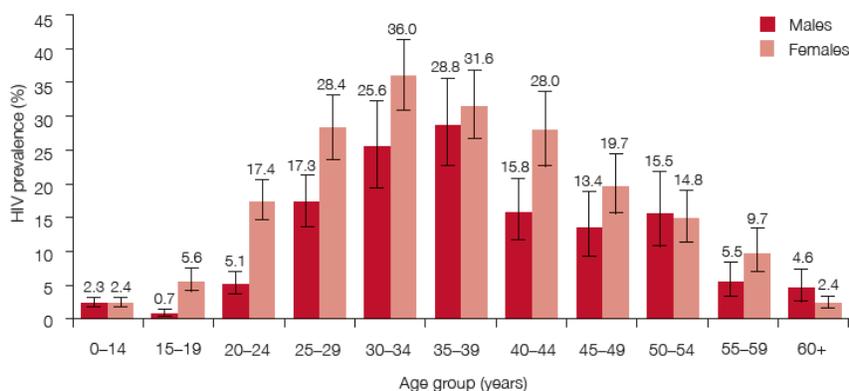


Figure 2. HIV Prevalence in South Africa by age group and gender (19).

In general, if men and women only chose sexual partners from their age groups, then the HIV in those populations would remain “trapped”, and as they grow older and die HIV would be eliminated (24). Although initially, the men in younger age groups would have increased risk of acquiring HIV from age-similar women (13).

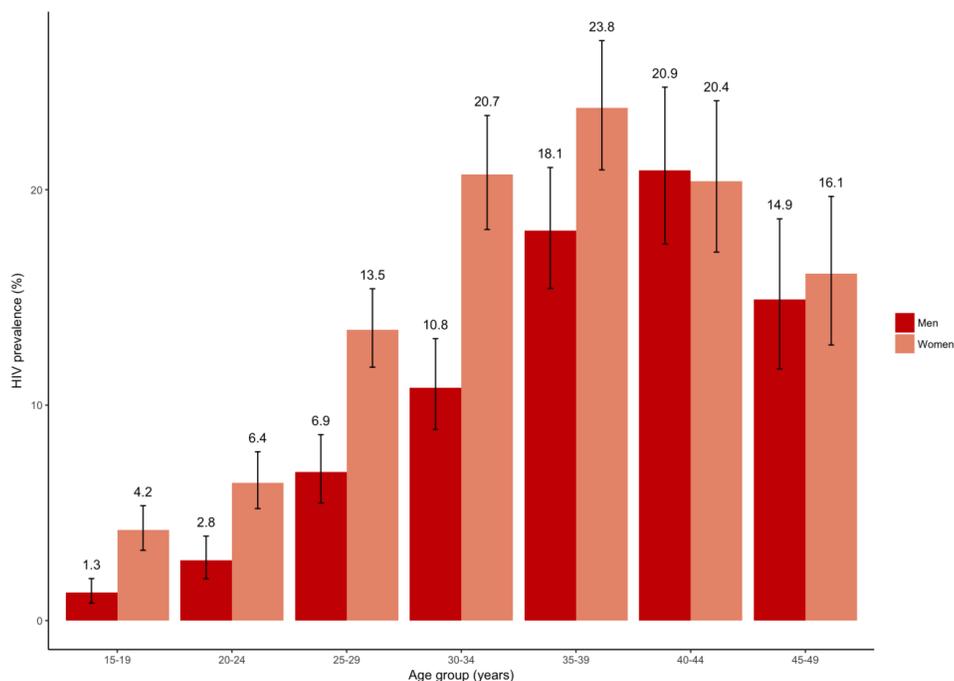


Figure 3. HIV Prevalence in Malawi by age group and gender, adapted from (20)

### 1.3.2 Sexual behaviour correlates of age-disparate relationships

In addition to the increased probability of choosing an HIV-infected male partner from older age groups, age-asymmetric relationships have also been independently associated with high-risk relationship characteristics. Power imbalances resulting in the inability of women to negotiate safer sex practices have been cited (14,25–28). One study indicated that young women may fear the reactions of their older male partners for suggesting using condoms during sex (29), while another qualitative study found that older men think they should not have to use a condom since they “bought the woman” (30). Indeed, several observational studies have found that women who are in age-asymmetric relationships rarely or inconsistently use condoms (29,31–36) and often report higher sex frequencies (29,31,33) than women in age-similar relationships. This implies that if a young woman enters an age-asymmetric, HIV sero-discordant

relationship where condom use is low and the frequency of sex is high, the per-partnership probability of transmission will also increase.

Age-asymmetric relationships have also been associated with older male partners engaging in multiple concurrent relationships (31–33,37,38). Concurrency is typically defined as having overlapping sexual relationships where sexual intercourse with one partner occurs between two acts of intercourse with another (39). One study of concurrency in the high HIV prevalence setting of Cape Town, South Africa found that 18% of the study population had concurrent relationships in the previous year, with the median duration of overlap being 7.5 weeks (40). There are other examples noted of high concurrency prevalence in high HIV prevalence settings (41–43). If we assume that an HIV-positive person's viral load peaks in infectiousness in the first couple of months after acquiring infection (44,45), this means that in populations such as those cited above where concurrency is prevalent and duration of overlap is long, a young woman in an age-asymmetric relationship may be at increased risk of acquiring HIV. In other words, concurrency will increase the likelihood that she will be sexually active with a man when he acquires acute HIV infection from an outside relationship. While concurrency of a male partner could hypothetically increase a young woman's risk of HIV infection, there is still conflicting evidence about a direct causal relationship between concurrency and HIV infection (46–54). Along these lines, there is evidence to suggest having multiple partners – concurrent or not – puts a person at increased risk of HIV infection (55–57).

There is also evidence to suggest that being in an age-asymmetric relationship may be related to alcohol use (29,32,58). Specifically, one study found that men who were in ADR with young women were more likely to drink before sex (32), while another found that binge-drinking was more frequent among girls whose partners were two or more years older (29). The relationship between alcohol use and HIV infection in sub-Saharan Africa has been well-documented and several reviews have reported a higher occurrence of HIV infection among people with problem drinking compared to those without problem drinking (59–61). Alcohol use in itself may influence sexual risk behaviour by limiting reasoning skills, judgement abilities, and personal responsibility for actions, while simultaneously increasing sexual arousal (62). This manifests as a person engaging in more casual (63,64) and unprotected sex (65,66).

Finally, there may be some evidence that young, unmarried women are having more anal sex than older women (67). The epithelium in the anus is relatively fragile compared to the vagina (68), and thus anal sex has been linked to higher risk of HIV transmission in heterosexual female sex (69,70). There is some qualitative evidence to suggest that anal sex may be linked to asymmetrical power relations between men and women in South Africa (71). However, the association between age differences and anal sex is understudied, with only one study showing a relationship between having anal sex without a condom in the previous 3 months and having a male partner that was two or more years older (29).

### **1.3.3 Biological susceptibility of young women to HIV infection**

As noted in Section 1.1, large age differences in sexual relationships have been proposed as an explanation for the large gender imbalance in HIV incidence among young adults in Southern Africa. However, it is important to point out that these may not be the only reasons for the gender imbalance in HIV incidence among young adults. There are

additional biological mechanisms which may make young women more susceptible to HIV infection.

It has been suggested that there are differences between men and women in the mucosal immune environment of the genitals (72). Women, for example, have a larger surface area at the site of HIV exposure (i.e. in the cervico-vaginal mucosa) than men (i.e. in the penis and foreskin). Moreover, it has been noted that semen can remain in the female genital tract for up to 3 days after a sexual episode (73). Compounding these effects, young women may have more gateways for HIV to enter their bloodstream due to increased susceptibility related to genital micro-abrasions. Before a woman undergoes her first pregnancy, her squamous epithelium lacks tight junctions and is more prone to micro-abrasions during intercourse (74). Though the epithelium may be damaged, the basement membrane remains intact, and thus, viruses may be able to bind to the exposed membranes (74). Female genital mucosa also tend to have more activated immune cells and express higher levels of the HIV co-receptor CCR5 than foreskin cells (72), implying that the mucosa of women are more likely to assist with HIV replication.

Herpes simplex virus 2 (HSV-2) and other STIs have been shown to be associated with HIV acquisition (75–78) owing in part to the STI's increased recruitment of HIV target cells to the mucosal site (79,80). A study conducted in four cities in Africa, found that HSV-2 was prevalent in the populations it studied, and at young ages women were more likely to be infected than their male counterparts (81). This suggests that if young women are more likely to be infected with HSV-2, their chances of HIV acquisition may also increase. Some studies have even found direct links between age-asymmetries in relationships and an STI diagnosis in young women (29,82,83).

Finally, the use of the injectable hormonal contraceptive (IHC), depot medroxyprogesterone acetate (DMPA), has also been shown to increase HIV acquisition in women (84,85). Indeed, Malawi and South Africa also have high prevalence of IHC use (86), which may partially explain the discrepancies between male and female HIV prevalence in these countries. In one study of hormonal contraceptive use in South Africa, there was a moderate effect of DMPA on HIV acquisition, with young women having a slightly higher hazard of acquiring HIV (87).

The increased biological susceptibility of younger women to HIV infection in combination with a tendency towards choosing older partners who engage in risky sexual behaviours may explain the stark contrast in HIV prevalences among young men and women.

### **1.3.4 Reasons for pursuing age-asymmetric relationships**

With all of the HIV-related risks that age-asymmetric relationships may pose for women, one may wonder why many still choose to engage in these relationships. Women typically cite economically motivated reasons, such as obtaining financial security (14,88–90), receiving pocket-change or gifts (25,91), attaining money for school fees (89), and becoming food-secure (92). There are also psycho-social benefits women describe, such as, increased self-esteem, love, emotional support, as well as improved status among peers (14,25,93). Indeed, one study of orphaned young women in Zimbabwe found that wives in ADR were better off in terms of the number of meals per day, number of assets, and quality of life measures (94). Another ethnographic study in Mozambique

found that women enjoyed these relationships because they could achieve independence within them (95), and felt as though they could be their own agents.

Young women cite many positive benefits of age-asymmetric relationships, but little is known about whether they perceive risks in these relationships. One qualitative investigation of cross-generational relationships in Kenya assessed women's perceptions of HIV and STI risk in these relationships (25). The young women in their study thought the greatest risk posed to them by these relationships was the possibility of being discovered by their own family, or the wife of the male partner (25). Importantly, the women in the study were not concerned about acquiring STIs because they judged their partners to be low risk (25). Perceptions of risk are context-specific. A more recent qualitative study conducted in Tanzania – after a radio campaign discouraging cross-generational relationships – found that cross-generational relationships were considered to be risky because older men who sought out young partners were thought to be deceitful, and that they would easily lie about their HIV status, as well as refuse to wear condoms (96). Similar views were articulated by community members interviewed in focus groups in Malawi, Botswana, and Mozambique (97). They thought that young girls were at risk of HIV infection primarily through ADR, which they associated with concurrency and less condom use (97).

#### 1.4 ASSOCIATION BETWEEN AGE DIFFERENCES AND HIV INFECTION

---

In Section 1.3, I proposed mechanisms for how age differences in relationships might facilitate HIV transmission. What is left unanswered is whether or not this relationship has been demonstrated empirically. In this section, I will critically evaluate the evidence that has been used to substantiate and, in some cases, contradict the hypothesis that age-mixing drives the HIV epidemic. Broadly speaking, the research that has been done falls into three types of study designs: modelling studies, ecological studies, and individual-level observational studies. The research falling into each of these designs will be considered in turn.

##### 1.4.1 Dynamical models

Before people studied age-mixing patterns in real-world populations, modelling studies established that there was a theoretical relationship between age-mixing and HIV transmission. In 1992 Anderson et al. published the results of their compartmental model that examined how age classes between two sexes, as well as age-dependent rate of sexual partner change, affected the spread of HIV in a population (10). They found that when the rate of sexual partner change would decrease with age or when males would have more sexual contact with younger women, these scenarios would magnify the epidemic (10). While this study was novel for the time, it was relatively simplistic and did not take into account heterogeneity of sexual activity within age and sex classes. Moreover, at the time of the study, little was known about HIV doubling times, rates of infection, and incubation periods for AIDS.

A few years later Garnett and Anderson published new results which incorporated heterogeneity of sex acts per partner and numbers of sexual partners (12). Their aim was not to provide quantitative projections of the epidemic, but rather to demonstrate how modelling heterogeneity in transmission dynamics may aide in interpretation of

epidemiological data. They confirmed that heterogeneity actually increases the likelihood of HIV persistence in the population, but in the long term may reduce the magnitude of the epidemic once it becomes endemic. Likewise, they established that assortative mixing (like with like) would foster rapid spread of HIV in highly sexually active people, but then result in a lower endemic equilibrium (12). In this study, assortative mixing was used in reference to the amount of sexual activity people had, not specifically to age-assortativity. From the perspective of my research, what was missing was how the magnitude and persistence of the HIV epidemic would be altered with different degrees of sex frequency and age-assortative or disassortative mixing.

Over a decade later, a mathematical model of the heterosexual HIV epidemic was developed by Hallett et al. to ascertain the population-level impact of reducing age differences among sexual partners (13). Specifically, they investigated what effect there would be of: 1. Changing the distribution of young women's partnerships with older men, while maintaining the total number of partnerships, and 2. Removing a certain percentage of partnerships for a given age-difference, resulting in less total numbers of partners (13). They found that if the number of relationships was preserved, the prevalence of HIV in young women would be reduced, while for men it would be temporarily increased (13). However, over a lifetime the risk of HIV infection was reduced by 10% for women and 5% for men (13). The effects were more drastic under the scenario of reducing the total number of partnerships. Both men and women would have immediately-reduced HIV prevalence, with their lifetime risk reduced by 25% for women and 22% for men (13). While the investigators considered variation in age differences, the nature of their deterministic model did not allow them to explicitly model relationships, and thus they could not distinguish between within-individual and between-individual variation of age differences.

More recently, d'Albis et al. conducted a modelling study, which investigated the effect of changing the distribution of age differences between sexual partners on the HIV epidemic (98). They were able to demonstrate that the average age-difference in a population has relatively little impact on the magnitude of the epidemic (98). Importantly, however, large variation in age differences, regardless of the mean age difference, had substantial impacts (98). While this study has crucial insights for studying HIV and age-mixing dynamics, it lacked realism and could have been improved by incorporating other information for determining how sexual contacts are made.

The EMOD-HIV model was used by Bershteyn et al. to identify key transmission events that allow HIV to sustain itself in younger generations and propagate efficiently, as well as to identify key population groups to target for transmission-blocking interventions (99). In their model, age determined the rate at which individuals entered casual and spousal relationships, the frequency of partnership formation, and the level of age-assortativity (99). The individuals who enabled transmission events, and were thus fit for interventions, tended to be STI co-infected males aged 26-29 and women aged 23-24 (99). The authors concede that their results are subject to the assumption of age-based heterogeneities and there are a large range of other assumptions they could have explored to identify subgroups for interventions.

#### **1.4.2 Ecological associations of age-asymmetry and HIV prevalence**

An article written by Wellings et al. in 2006, attempted to describe patterns of sexual behaviour in 59 countries for which data were available (100). Of importance to my

research was the finding that out of all the regions in the world, countries in the African continent tended to have the largest mean age-differences (100). The implication here was that this population-level observation may also explain why sub-Saharan Africa has such a high prevalence of HIV.

Another ecological study compared HIV risk factors in adolescents from two high HIV prevalence (South Africa and Zimbabwe) and two low HIV prevalence settings (Tanzania and Uganda) (11). They observed that the prevalence of women with older partners was higher in high-HIV prevalence settings (11). One other ecological study taking place within Kenya found that there was no association between the proportion of women in ADRs and HIV prevalence by ethnic group (101).

While two of the studies here lend support for the hypothesis that age differences in relationships are related to HIV transmission, they can only be considered hypothesis-generating at best, and say nothing about whether individuals are at increased risk of acquiring HIV infection when they are engaged in age asymmetric relationships. Studies of this nature often suffer from ecological fallacy. In other words, they attempt to make an inference about the nature of individuals based upon these large population-level aggregate statistics (102).

### **1.4.3 Individual-level observational studies of age-asymmetry and HIV infection**

In addition to population-level insights that have been garnered through dynamical models and ecological studies, over the past 15 years new understandings of the risk that age differences pose for HIV transmission have been developed through individual-level observational studies.

A household survey taking place among 15-54 year old men and women in rural Zimbabwe in 2000 was one of the first studies to demonstrate an individual-level association between age differences in sexual relationships and the odds of a participant being HIV positive (31). Several cross-sectional, nationally-representative surveys that took place in South Africa in 2003 (23), as well as 2002, 2005, 2008, and 2012 (103) also demonstrated a positive association between young women being in ADR and being HIV infected. Kaiser et al. analysed couples' data from a nationally-representative survey in Kenya and found that when male partners were 10 or more years older, the odds of the partnership being HIV-discordant was 1.6 times greater than when partners were less than five years older (104). Unfortunately, it is unclear whether or not it was the older male partner who tended to be HIV infected, thus putting his younger female partner at increased risk of HIV infection.

While the findings from these studies are useful and interesting, there are two important limitations of them. First, there may be social-desirability bias due to the face-to-face interviewing methods the studies used. Particularly with sexual behaviours the bias tends to be gendered, with men over-reporting and women under-reporting sensitive sexual behaviours such as the number of sexual partners they had (105,106). Unfortunately, when this bias is present it is difficult to determine whether the measures of association would be biased away from or towards the null. Second, both studies are cross-sectional and therefore temporality between the exposures and outcomes cannot be established (18). In other words, we do not know whether a person acquired HIV as the result of being in an ADR, or whether they acquired HIV first from

engaging in some other risky behaviours and then subsequently began an ADR after infection occurred.

For some studies investigating this relationship, temporality was not an issue because the studies were longitudinal and measured HIV incidence. One such study used data from a randomised community trial for AIDS prevention in rural Rakai District, Uganda between 1994 and 1998 (107). The investigators found that among women who were 25-29 years old, having a partner who was 10 or more years older was associated with an 80% reduction in risk of acquiring HIV over the period of observation (107). Among women in other age groups and with other partner age differences, there was no increased or decreased risk of becoming infected with HIV (107). A more recent, high-profile study of a large population-based cohort in rural KwaZulu-Natal, South Africa was not able to find evidence of a relationship between age differences and HIV infection among women aged 15-30 (108). However, in the same population among women older than 30 years, HIV incidence decreased as partner age differences increased (109). Two other longitudinal studies taking place in South Africa also could not find strong evidence of a relationship between ADRs and HIV incidence (38,110). Only one recent longitudinal study in Zimbabwe, of women aged 15-24, found a positive association with increasing partner age differences and HIV incidence (111).

Like the cross-sectional studies reported above, these longitudinal studies also used face-to-face interviewing (FTFI) methods, and were thus likely to suffer from the same forms of bias. In addition to this, all of the studies described so far, including the longitudinal studies, only measured age differences in the most recent sexual relationship. This is an important limitation because it does not consider the age differences in other relationships that may have occurred, and which may be the relationships where the HIV transmission event took place.

A recent cross-sectional survey of 15-49 year olds in two different KwaZulu-Natal, South African populations was able to get around this limitation by conducting a phylogenetic study which identified clusters of linked infections (112). They concluded that this population is characterized by cycles of HIV transmission where girls 15-25-year-old are probably infected by men in their 30's, on average. Those same men were likely to have acquired HIV from women aged 25-40 (112). This study indicates that there is, indeed, age-disassortative mixing happening for sexual partnering in this high HIV prevalence population. Unfortunately, studies of this nature are difficult to do because of the prohibitive costs of gene sequencing for large population based samples.

One theory for why most of the cited longitudinal studies did not find a relationship between age differences and HIV is that the older partners of those women may have been more likely to be on ART (113). The authors of an interesting recent cross-sectional survey in South Africa specifically investigated whether men in ADR were more likely to be HIV positive and whether ART mitigated that association (114). They found that men with a 5-9 year age difference and 10+ year age difference in relationships with women 15-24 years old were more likely to be HIV positive, and the 5-9 year age differences were also associated with the men being HIV positive and ART-naive (114). Men in relationships with 25-29 year olds having 10 or more years age difference had a reduction in odds of being HIV positive and ART-naive (114). These results suggest that the relationship between ADR and HIV may change across context according to ART coverage. One key limitation of this study in addition to the ones noted above for cross-sectional surveys is that they frame their findings in terms of the risk the ADR pose for

their female partners. However, they did not know the status of the female partner, and thus they could not exclude the possibility that the men received HIV from their partners. Indeed, if we re-examine the results from the previously cited phylogenetic study, we see that it is entirely possible that the men actually got infected by those partners.

## 1.5 MOTIVATION

---

The studies highlighted above demonstrate that HIV research has come a long way over the past 25 years towards describing and identifying sexual behaviours that may put young women at increased risk of infection. However, the evidence for the hypothesis that age-asymmetries facilitate HIV infection is equivocal. It may be the case that in some contexts age differences do not predict HIV infection because the age-profile of HIV prevalence among men changes with time and location, and thus young women who choose older male partners from an age cohort with a low HIV prevalence may not be at increased risk (e.g. Kelly et al. (107)). Moreover, in other settings there might not be sharp wealth gradients between women and men who are in ADR, and thus less power imbalances (e.g. Harling et al. (108)). And yet, in other contexts ART coverage may mitigate HIV transmission in ADR (e.g. Evans et al. (114)). The reality is that the scientific community still does not know in which contexts age differences may be important for HIV transmission. Different milieus will drive individuals to adopt different informal rules for choosing partners (115,116). To this end, in three of my manuscripts that are featured here I have described and visualized age-mixing patterns in different populations from South Africa and Malawi (See Chapters 3, 5 and 7).

The literature review also revealed that there are some crucial methodological limitations to the way age differences have been studied in relation to HIV infection. For one, most studies, including the cohort studies and randomised controlled trials cited earlier, only reported on age difference measures from the most recent relationship reported by participants. Importantly, in order to establish a causal relationship between age-differences and HIV infection, one needs to know which relationship was likely to have caused the transmission event (112). Alternatively, in lieu of phylogenetic data, one could improve the study of this association by examining other indicators of age-mixing which may be better predictors of cumulative risk across relationships. For example, the bridge width may be more meaningful. In Chapters 5 and 7, I explore the relationship between these new measures of age difference and prevalent HIV infection, as well as discuss their importance, utility and limitations.

The studies that use the most recent partner age difference as their primary exposure of interest often justify this choice by the fact that the vast majority of participants only report one sexual partner. I would argue that if they are only reporting one partner, a likely cause would be the mode of interview, rather than the participant actually having only one partner. As previously mentioned, FTFI often results in social desirability bias, and one of the most common ways this manifests itself is in reporting only one partner (117,118). One way to prevent or reduce this bias might be to use Audio Computer-Assisted Self-Interviewing (ACASI) (118–121). In a previous study we conducted, we found that ACASI administered in a Cape Town sexual behaviour survey, may have produced more accurate results by removing some of the gendered social desirability bias (100, Appendix A). Additionally, it elicited more reporting of sexual behaviours (e.g. number of sexual partners) compared to some Demographic and Health Surveys (100,

Appendix A). Two of the studies reported in this dissertation, use ACASI for the mode of interview (See Chapters 5 and 7).

A final limitation of previous age difference studies is that most do not take into account the role that women may play in transmitting HIV. We now know from the previously cited phylogenetic study that young women are not just passive recipients of HIV, but that they, themselves, are often the transmitters (112). Women are also agents in their sex lives, making complex risk decisions and trade-offs which may knowingly or unknowingly put themselves and men at risk (See Chapter 4). Therefore, simplistic analyses that measure the risk of HIV infection for women in relationships with older men may only be telling part of the “age-mixing story”. What is needed is an analytical paradigm shift that allows for more complex modelling of age, age differences, and age difference variability (both between- and within-individual variation) in relation to HIV transmission in both women *and men*. These dynamics will also be explored in Chapters 5 and 7.

Understanding age-mixing will ultimately be important for identifying specific groups at risk of acquiring or transmitting HIV, and thus for targeting HIV risk reduction interventions. Moreover, having a keen awareness of the sexual behaviours that take place within age-asymmetric relationships (See Chapters 3, 5, and 6), as well as psychosocial factors which may prompt women to begin them (See Chapter 3 and 4) will allow researchers and other stakeholders to fine-tune those interventions for maximal efficacy.

## 1.6 OBJECTIVES

---

In order to assess whether age-asymmetries in sexual relationships have the potential to sustain the HIV epidemic in Southern Africa and increase individual-level risk of acquiring HIV, I aimed to:

- 1 Describe and visualize age-mixing patterns in different populations of South Africa and Malawi;
- 2 Identify socio-demographic predictors and psycho-social motivations for engaging in relationships with older men; and,
- 3 Ascertain whether age differences in relationships are correlated with an individual’s HIV infection status, or other sexual risk behaviours commonly associated with increased HIV transmission risk.

## 1.7 THESIS STRUCTURE

---

In the next chapter I will expand upon the methods of my research, providing a synopsis of the datasets I used, as well as common statistical methods that featured in many of my manuscripts (Chapter 2). Chapters 3 to 7 are where readers will find the collection of the manuscripts that feature the main results of my dissertation research. They are ordered chronologically, and each of the results chapters will begin with the research context that prompted us to carry out the specific investigation presented in the chapter. In Chapter 8, I will discuss the main results emanating from this work and highlight their implications for further research. Finally, in Chapter 9, I will make some closing remarks and recommendations for future public health interventions that are based upon the findings presented here.

## CHAPTER 2: METHODS

---



Photo credit: Alexis Strimenos

## 2.1 STUDY DESIGN AND OVERVIEW

---

The results I present in Chapters 3-7, come from four different secondary statistical analyses and one qualitative study. All of the secondary analyses, with the exception of one (See Section 2.2.4) use data that were collected by other institutions. In this chapter, I provide an overview of the source of the datasets used in the analyses, outlining details about the study settings, innovative features of the design, and what was measured for each one. After which, I describe my own methods for analysis. Without giving too many details – which can be found in the methods section of the manuscripts presented in Chapters 3-7 – I will give a quick summary of the most commonly used statistical methods that feature in my analyses. Finally, I will end the chapter by supplying the details for ethics approval and how the results have been disseminated.

## 2.2 AN ABRIDGED DESCRIPTION OF THE DATASETS

---

### 2.2.1 Cape Area Panel Study (CAPS)

#### 2.2.1.1 *Study synopsis*

The Cape Area Panel Study (CAPS) collected the data that feature in Chapter 3. It was a cohort study carried out by the Population Studies Center in the Institute for Social Research at the University of Michigan and the Centre for Social Science Research at the University of Cape Town. In a series of waves, it investigated the lives of nearly 4800 adolescents and young adults aged 14-22 years old living in Cape Town, South Africa: Wave 1 (2002), Wave 2a (2003), Wave 2b (2004), Wave 3 (2005), Wave 4 (2006), and Wave 5 (2009) (122).

The 4750 participants participating in Wave 1 were selected from a three-stage sampling procedure. First, census-demarcated Enumeration Areas (EAs) were stratified according to the predominate race present in the EA, and then they were randomly sampled within each stratum with a probability proportional to the population size. Second, households were randomly sampled within the previously selected EAs and kept for final inclusion if they had family members aged 14-22 ( $n = 5256$ ). Finally, within the households up to three participants aged 14-22 were selected (122).

We analysed data from 1960 young people who participated in Wave 4, who had previously had sex, but were never married. In the survey from this wave, questionnaires were administered and they solicited responses about schooling, relationships, employment, family and support systems, as well as sexual and reproductive health issues (122).

#### 2.2.1.2 *Study setting*

Cape Town is the second largest city in South Africa, with a population of approximately 4 million people living in 1.2 million households (123). In South Africa racial classifications are important determinants of health outcomes and socio-economic opportunities, due to the country's previous history of Apartheid (124). The 2011 census indicated that a plurality of the population was coloured (42%), followed by black (39%), white (16%) and other racial groups (3%) (125). The term 'coloured' has a specific definition in the South African context and it is used to refer to racially-mixed descendants of European, indigenous, and/or Asian slave origins.

A large fraction of the Cape Town population is impoverished and lacks secondary education. In 2016 the proportion of the population that was below the poverty line was 39.3% and the percentage of households with no income was 13.9% (123). While most people in Cape Town reside in formal dwellings, 18.6% live in “shacks” (123). As of 2011 only 46% of those older than 20 years had completed grade 12 or higher (125). Only 33% of City of Cape Town grade 10 students in 2014 enrolled in grade 12 in 2016 (123).

Cape Town is also characterized as having a large number of health and safety concerns. In 2016 the murder rate was 62 per 100,000 and the sexual offense rate was 104 per 100,000 (123). Tuberculosis is another major health concern in the city, with over 24,000 patients being treated for TB in 2016 (123). The background HIV prevalence for the City of Cape Town in 2012 was 5.2% (95% Confidence Interval (CI): 3.4% – 7.8%) (19) and 13.0% of premature deaths were caused by AIDS (126).

## **2.2.2 Likoma Network Study (LNS)**

### *2.2.2.1 Study synopsis*

The Likoma Network Study (LNS) was a cohort study investigating sexual behaviour on Likoma Island, in Lake Malawi, Malawi. This study began in 2005 and was originally carried out by researchers at the University of Pennsylvania and the University of Malawi College of Medicine. At the time I began my analyses, the LNS had completed two rounds of data collection, taking place in 2005/2006 and 2007/2008, with a third round slated for 2013.

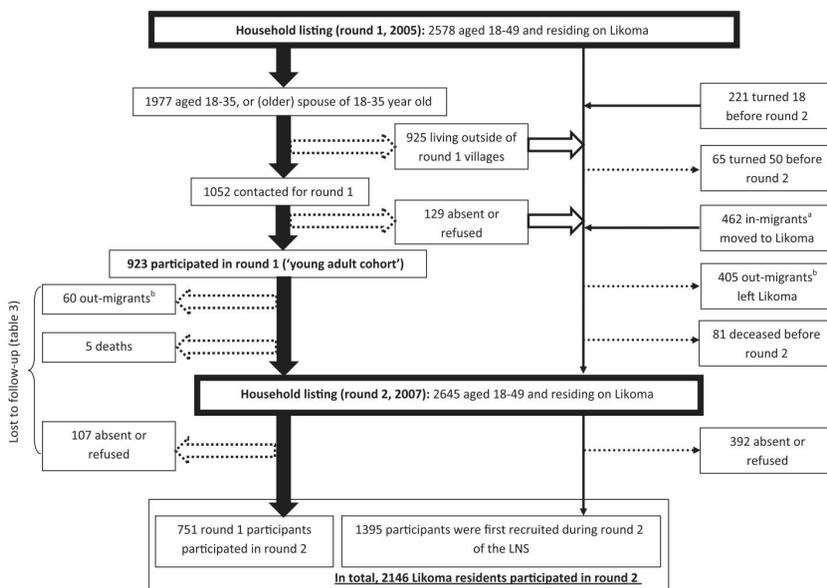
This study uniquely featured a socio-centric study design in order to map out all of the sexual connections within this network (127). First, a census of all island residents and households was conducted in order to obtain a complete list of potential network members. Second, a sexual behaviour survey was administered to all eligible adults on the island, during which they could nominate their five most recent sexual partners. Finally, the researchers attempted to anonymously link those partners to the list of residents on the island (128). This process did not involve contact tracing: all people who participated were contacted because they lived on the island, not because someone named them as a sexual partner (127).

In the first round (2005/2006) 2578 residents were listed in the household roster. However, the survey was only administered to residents in seven of the 11 island villages, and to those who were between the ages of 18 and 35 years or were married to someone in that age range ( $n = 923$ ) (127). In the second round of data collection (2007/2008) 2645 residents were present for the household listing. In this round, during the survey phase, all villages were included and those between 35 and 49 years could also participate in the survey (127). This resulted in 2146 participants taking the survey with 1395 taking the survey for the first time. Figure 4 shows the LNS participant flow chart (127).

The survey questionnaire was administered using ACASI, in order to create a sense of privacy and anonymity. While speaking into a recording headset, participants could list their partners, and identifying characteristics such as the partner’s location within the village and his/her nicknames. In addition to nominating sexual partners, participants answered questions about frequency of sex, condom use, and other relationship attributes (127).

The HIV status of participants was determined by means of rapid HIV tests. Health councillors administered the tests and provided counselling in participants' homes. Unigold (Trinity Biotech, Ireland) and Determine (Abbott, Japan) tests were run concurrently and tests that were in agreement were considered to be a true result. In the event of conflicting test results two additional tests were administered. If all tests were discrepant, then the participant was referred to a local centre for further testing (127,128).

In the first round of data collection partner ages were not available, and therefore we conducted our analysis with only second round data. Thus, our analysis in Chapter 5 includes 1922 participants from 2007/2008 who reported at least one sexual partner in the previous three years. Together, they provided details on 3336 relationships.



**Figure 4. Likoma Network Study Participant Flow Chart. Extracted From Helleringer et al. 2013 (127)**

### 2.2.2.2 Study setting

Likoma is a small, isolated island in the eastern part of Lake Malawi that was considered to be a high-HIV prevalence setting in 2005 (128,129). Chizumulu, another small island, is located to the West of it, while the Mozambican shore is located to the East. Likoma, itself, was home to about 7,000 people who primarily identified as Nyanja (75%), Chewa (10%) or Tonga (10%) ethnic groups. Over 50% of the residents were

under the age of 15 (129) at the start of the study. Education was widespread on the island, and most young people stayed in school until their late teenage years (129).

Residents of Likoma tended to delay marriage until their schooling was done. When a marriage occurred, a woman would move into the household of her new husband's family, which was usually located in a neighbouring village. Polygamy was not uncommon, nor was divorce on the island (129,130).

At the time of the study, the local island economy was under-developed and the primary form of income for most households was through fishing-related activities (128). Though Likoma was hard to reach – by ferry it took 7-20 hours from mainland Malawi –, residents would come and go throughout the year primarily for economic activities. From April to November they would often travel to Mbenje Island in the South in order to fish and throughout the rest of the year they would make trips to mainland Malawi to sell their fish or take part in trade (127).

The island had several key features which made it the perfect place to study HIV transmission dynamics through sexual networks. First, the island was fairly remote and travel to and from it was not easy. This made it possible to identify many sexual partners, since the population was somewhat bounded. Second, the population was relatively small, so it was logistically feasible to identify and survey most sexually active residents on the island. These two factors coupled with the relatively high HIV prevalence on the island, made it the perfect “laboratory” to identify aspects of sexual mixing that could contribute to the spread of HIV.

## **2.2.3 Schooling, Income and Health Risk (SIHR) Study**

### *2.2.3.1 Study synopsis*

The Schooling, Income and Health Risk (SIHR) study was a cluster randomised trial of a cash-transfer experiment administered to girls in Zomba district, Malawi (131). The original study was conducted by researchers at The World Bank Development Research Group, George Washington University, the University of California San Diego, and the University of Malawi. The dataset resulting from this study was used in Chapter 6.

Census EAs in the district were randomly sampled ( $n = 176$ ) from urban, rural, and near-urban areas. Within each sampled EA, all households were listed and a representative from each household was asked if there were any never-married girls in the household who were between the ages of 13 and 22. Households that had at least one girl who met the criteria were included in the study (131).

Half the EAs were randomised to the intervention group ( $n=88$ ), and the others to the control group ( $n=88$ ). The control and intervention group EAs included both schoolgirls and dropouts at baseline (131). The intervention schoolgirl EAs were further stratified into those who received cash transfers unconditionally and those who received cash transfers only if they attended school at least 80% of the days school was in session. Some of the girls in both of the intervention EA groups did not receive any cash transfers. For more details see Figure 5, which contains the previously published trial profile (131). Cash was administered to both households (\$4-10 per month) and girls (\$1-5 per month) in the intervention groups (131).

A survey was administered to the girls on four occasions throughout the study, though at the time I conducted my analysis, there were data publicly available for only three rounds (132). Round 1 (baseline, 2007/2008) took place before allocation of intervention assignment (133). Round 2 (R2, 2008/2009) surveys were held twelve months after the intervention assignment (133). Round 3 (R3, 2010) happened two years after the intervention assignment and one year after the end of the intervention (133).



Figure 5. Schooling, Income and Health Risk Study Enrolment and Participation Flow Chart. Extracted from Baird et al. 2012 (131)

The first part of the survey was administered to the head of household and obtained information on household characteristics. The second part was administered to the girls who took part in the study and it focused on health, sexual relationships and friendships. At 18 months – between R2 and R3 – HIV and HSV-2 biomarkers were collected, though this data was not publicly available (132).

The study researchers found that the prevalence of HIV and HSV-2 was lower among the girls in the intervention arm than the control group at their 18-month follow-up (131). Importantly, and of interest to my research, they also found that the intervention groups had a lower prevalence of engaging in relationships with men 25 years or older (131). Thus, we hypothesized that the choice of younger partners by those in the intervention group may partially explain the lower STI prevalences that were observed.

A total of 3976 girls were sampled and interviewed at baseline, of which 2,907 were schoolgirls and 1069 were dropouts. Like the LNS, ages of partners were not collected at baseline. Therefore, I focused my analysis on the 1108 schoolgirls who reported a relationship in R2 or R3 (relationships n = 1491).

### *2.2.3.2 Study setting*

Zomba district is in southern Malawi and it includes one of the four largest cities in Malawi (Zomba City), in addition to a large rural population (134). Zomba city had a population of approximately 88,000 in 2008, and at that time it was expected to have 3% annual growth (135). It was primarily agricultural, with most people participating in subsistence farming (136), and only 6% of the adult population having a formal income as of 2008 (134). It was noted for being the third poorest district in Malawi in 2009 (134), with over 60% of the population living in informal settlements (135). In 2008 the per-capita exchange rate comparable consumption was only 20.60 USD per month (134).

In contrast to the LNS population, most students did not complete their secondary education and 28% of women, and 24% of men were illiterate (135). There was a shortage of secondary schools in Zomba, with most having inadequate equipment and an uneven distribution of teachers (135).

Like the schools, the hospitals and clinics in Zomba were also under-staffed and lacking proper medical equipment (135). In addition to using hospitals Zomba residents also saw traditional medicine practitioners and traditional birth attendants for health-related services (135). HIV/AIDS was, and still is, one of the largest public health problems facing the district (135). As of 2008, the prevalence was nearly 14%, with most infections occurring in women (135).

## **2.2.4 Cape Town Sexual Behaviour Study (CTSBS)**

The Cape Town Sexual Behaviour Study (CTSBS) was a study carried out by my supervisors and I, in collaboration with researchers at the South African Centre for Epidemiological Modelling & Analysis and The Desmond Tutu TB Centre, both at Stellenbosch University. The study took place in three urban communities of Cape Town, South Africa: Delft, Wallacedene, and Khayelitsha. The study consisted of a sexual behaviour survey (featured in Chapter 7 and Appendices A-D) and a qualitative study of sexual relationship dynamics (Chapter 4).

### *2.2.4.1 Synopsis of the sexual behaviour survey*

The sexual behaviour survey component of the CTSBS was cross-sectional and data collection took place from July 2011 to February 2012. In order to gain a representative sample of the different races most affected by HIV in South Africa, we chose one predominantly black community and two racially diverse communities, consisting of black and coloured residents.

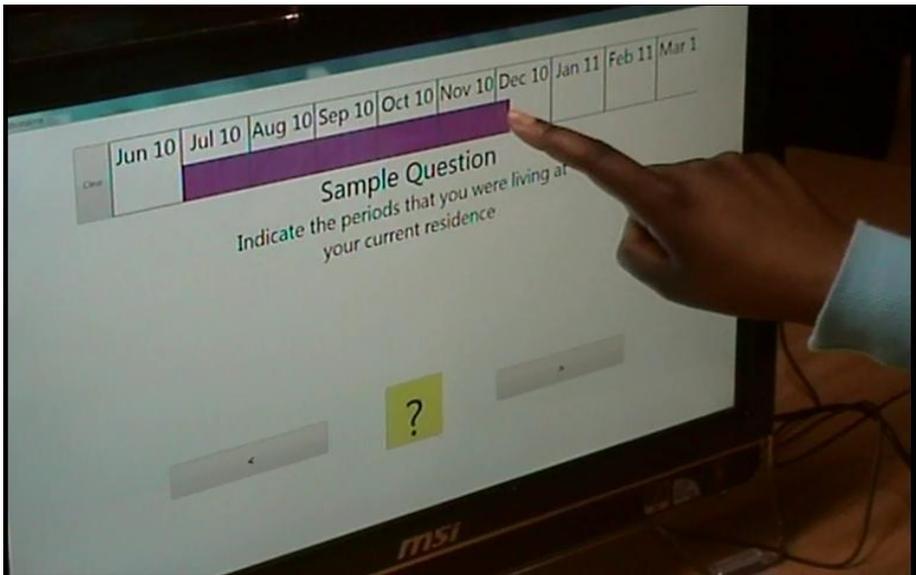
We sampled participants from another study, the Zambia South Africa TB and AIDS Reduction study (ZAMSTAR). ZAMSTAR was a community-randomised trial that aimed to reduce the prevalence of tuberculosis (TB) in communities with a high burden of TB and HIV, by using novel public health interventions (137). The ZAMSTAR study also provided HIV tests and counselling for willing participants. Our sampling frame consisted of 1857 randomly selected previous ZAMSTAR participants in the three chosen communities. We were able to locate 1115 (60.0% contact rate), and of those, only 1028 were still eligible to take the survey. In the end, 878 participants took the survey (85.4% response rate) and reported 1128 relationships.

A touch screen questionnaire, utilizing an ACASI application was developed to facilitate the collection of relationship history data (See Figure 6). The participants wore headphones to hear questions in their choice of English, Afrikaans, or isiXhosa and they selected their answers on a 22-inch touch screen monitor. If literate, they could simultaneously read the questions in the chosen language.

We collected detailed, retrospective sexual histories for a maximum of five main partners and 15 casual sexual partners. The questionnaire began by asking the participant basic demographic information and then proceeded to ask the participants about their main partnerships. Participants could then indicate, on a touch screen timeline, the periods they were in this relationship. For each relationship, we asked what year their partner was born (alternatively, they could indicate the current age of that partner), in addition to other partnership-level characteristics. For each distinct period that the participant engaged in sexual activity with a particular partner, measured in quarter-month intervals, additional questions were asked about the frequency of intercourse and condom use in that episode. This series of questions was asked for each main sexual partner, as well as the casual sexual partners. The duration of each relationship episode was displayed on the touch screen timeline for the participant to see, using different colours for each partner.

In addition to the sexual behaviour questionnaire we added three questions to assess the ease of use, perceived confidentiality and self-reported truthfulness when answering questions in this survey. Finally, we asked them to indicate their preferred mode for answering questions about sexual behaviour: touch screen computer with ACASI, researcher-administered verbal questionnaire, self-administered written questionnaire, or telephonic survey.

After the survey was complete, we anonymously linked the participant survey data to their ZAMSTAR HIV test results using a unique ID. In the ZAMSTAR survey, some participants chose not to have HIV testing done. For this reason, a small number of our random sample per study site consisted of participants who had a missing HIV status. We published additional details of the sexual behaviour survey elsewhere (See Appendices A-D). While I was a co-author on those manuscripts (the lead author on two), the results from them do not help to achieve the objectives of my PhD studies, and thus are not included in the results chapters of this dissertation (40,117,138,139).



**Figure 6. Demonstration of ACASI on touchscreen computer**

#### *2.2.4.2 Synopsis of the qualitative Interviews*

After completion of the sexual behaviour survey, I conducted in-depth, open-ended interviews with 23 women living in either Delft, Wallacedene, or Khayelitsha, Cape Town. We decided to undertake this ancillary study in order to aid our understanding of why women engage in risky sexual behaviours in this study setting, as well as to explore their perceptions and decision-making strategies in regards to their sexual partnerships.

I took a convenience sample of residents, that were not part of the sexual behaviour survey. They were recruited using a combination of maximum variation and snowball sampling. To be eligible, they had to be between the ages of 18 and 65 years old. The women were approached near the community town hall, and asked if they were interested in being interviewed once the purpose of the research was explained to them. After the interview, the women were then asked if they knew other women in the community who might be interested in participating in the study. Those women were approached in a similar manner. I stopped conducting interviews in a particular community when no new concepts emerged in the interviews. In total, 13 black and 10 coloured women, aged 20-59 years old partook in the interviews.

The interviews were conducted in the participant's home – if no family members were present – or in an unused room of the community town hall. A multilingual research assistant helped me to administer the interviews. They lasted approximately 30 to 90 minutes and were conducted in the language of the participant's choice; either English, isiXhosa, or Afrikaans. A thematic question guide, was used to conduct the interview. It contained open-ended questions, related to broad themes like, “age-disparate

relationships”, “multiple concurrent partners”, and “condom use”. Definitions were provided for some of the academic and scientific terms used in the questioning. The topics in the question guide were loosely adhered to, and I often changed the scope of questions as meaningful and unanticipated themes emerged. The participants also had the ability to change the direction of the interview. All interviews were digitally recorded, transcribed verbatim, and then translated to English.

#### *2.2.4.3 Study setting*

Delft, Wallacedene, and Khayelitsha are three disadvantaged communities of Cape Town, South Africa with a high prevalence of HIV. During Apartheid, in the 1980’s, all three locations were designated as “townships” – places where black and coloured South Africans were legally allowed to live. Khayelitsha was specifically built to house the large number of black migrants that were moving from other provinces in South Africa, as well as the overcrowding in other Cape Town townships (140). In the late 1980’s Delft was established and became the first mixed race township in Cape Town, which allowed black and coloured residents to live together (141). Unlike Khayelitsha and Delft, Wallacedene was not planned, and to this day it largely remains an informal settlement (142).

Khayelitsha has the largest population of the three townships. In 2011 it had 391,749 people, of which 98.6% were classified as “black” (143). Delft’s population was approximately 150,000, and it had roughly equal numbers of black and coloured residents (144). Wallacedene had the smallest population (approximately 10,000 people) and three-quarters were black, and the remaining quarter coloured or other racial groups (145). In a census that was done in 2001, all three communities had greater than 40% unemployment and under 25% of the population had completed secondary schooling (146). Much of what was expounded on in Section 2.2.1.2 applies to these three communities, though specific and recent statistics for each were hard to find.

### **2.3 FREQUENTLY USED ANALYTICAL METHODS**

---

Below I provide a brief non-technical synopsis of the statistical and qualitative methods I used to analyse my data. The overview discusses the techniques in general terms and provides the rationale for their use in particular instances. For detailed explanations of how the different methods were jointly used to provide results for specific research objectives, please see the manuscripts in Chapters 3-7.

#### **2.3.1 Regression modelling**

Each of the secondary statistical analyses presented in the results chapters features regression modelling. Regression modelling refers to statistical methods that estimate the relationship between a dependent variable (i.e. outcome of interest) and independent variables (i.e. exposure of interest, or predictors). The estimates produced by the models, help researchers to understand how the outcome variable will be affected, when any particular predictor variable is changed, and the other covariates are held constant. The regression modelling techniques described below are commonly used in statistical analyses today. Readers who want more details on how to perform the analyses, conduct diagnostics tests, or other technical details can consult e.g. Harrell 2015 (147). A summary of the regression modelling strategies I used are presented in Table 2.

In Chapters 5, 6, and 7, I used linear regression models to estimate the effect of different independent variables (e.g. participant age, study group, etc.) on age differences in relationships. Linear regression is typically used when the outcome of interest is continuously distributed and one expects the outcome to be a linear combination of the other predictors in the model. Through a series of algorithms (e.g. generalised least squares)  $\hat{\beta}$ -coefficients are produced by the models, which represent the slopes of the lines that fit the data points the best. As an example, the  $\hat{\beta}$ -coefficients can be interpreted as the average amount of change you would expect in partner ages for a 1-unit increase in participant age. Standard errors (SEs) of the  $\hat{\beta}$ -coefficient are also calculated by the models. These provide an indication of how far the sample  $\hat{\beta}$ -coefficient is likely to be from the true population  $\beta$ -coefficient. In the manuscripts, I have presented 95% Confidence Intervals (95% CIs). The upper limit of the 95% CI is roughly the  $\hat{\beta}$ -coefficient plus two times the estimated SE, while the lower limit is approximately the  $\hat{\beta}$ -coefficient minus two times the estimated SE. Finally, in a linear regression model the y-intercept of the best fitted line is also calculated by the model, and this equals the average response when all the covariates are at their baseline – reference – value.

Linear regression models can be extended to include other forms of outcome variables. This framework for modelling discrete or bounded dependent variables is referred to as generalised linear models (GLMs). In Chapters 3 and 5 we wanted to estimate the effect of different predictors on the probability of being in an ADR and being HIV positive, respectively. Each of those outcomes were binary (e.g. HIV negative status = 0 and HIV positive status = 1), and thus we used logistic regression. Logistic regression is a GLM which uses a link function – in this case, the logit function – to transform the linear combination of the independent variables into a function that predicts the odds of being HIV positive. In general link functions make the expected value of the outcome variable linear and the variances homogeneous. If the  $\hat{\beta}$ -coefficients produced by logistic regression are exponentiated, they can be interpreted as ORs (See Section 1.2.2). As an example, if an OR for an age difference predictor equals three, we might interpret it in this way: a 1-year increase in age difference increases the odds of being HIV positive three-fold.

Another type of GLM is a negative-binomial model. In some of the models described in Chapters 5 and 7 we wanted to determine whether certain characteristics were predictive of bridge widths (See Section 1.2.1). Bridge width is a variable that has a negative binomial distribution. In other words, it takes on discrete, non-negative values which are over-dispersed (the variance is larger than the mean). These models predict the expected count of the outcome variable, given a set of predictors. Like the logistic regression models, the  $\hat{\beta}$ -coefficients can be exponentiated to produce a more interpretable measure of association. In the bridge width models, the exponentiated  $\hat{\beta}$ -coefficients represent the ratio of expected bridge widths in one category of the predictor compared to the other category (Expected Bridge Width Ratio, EBWR). For instance, if the EBWR for HIV status is three, we could say that the expected bridge width for those infected with HIV is three times that of those who are HIV negative.

When the outcome of interest is an ordered categorical variable, the form of GLM that is used is called ordinal regression. Ordinal regression models predict the cumulative probability that an observation will fall into a particular category of the outcome or below. When exponentiating the  $\hat{\beta}$ -coefficient produced by the model, a POR is produced

(See Section 1.2.2). For example, in Chapter 6 we use ordinal models to predict the degree of condom use (Never/Inconsistently/Always) for given covariates. One way you could interpret a POR that equals three for age difference is: each year increase in age difference was associated with 3 times the odds of having more consistent condom use.

A final class of regression models I used is the Cox Proportional Hazard Model (CPHM). This is just one type of model that can be used to link survival time to model predictors. “Survival” is a general term that is used to describe the time to an event of interest. A participant is considered to have a “censored” outcome when they do not experience the event of interest during the period of observation. Censoring can result from either the study ending before the event occurred, or the participant drops out of the study and the event was not experienced before dropping out. CPHM can handle censoring though use all available information. CPHMs do not make any assumptions about the distribution of survival times. They do, however, assume that the effect of the covariates on experiencing the outcome is constant over time. In Chapter 6, we were interested in determining if age differences among the girls were related to the duration of their relationships. In this example the duration was censored, because we did not observe the end of the relationship among participants who indicated their relationship was ongoing. Put in terms of CPHMs, we wanted to know if the age differences predicted the hazard of ending a relationship. Once again, by exponentiating the estimate from the model, it was possible to produce a HR (See Section 1.2.2). If the HR was above 1.0, it indicated that the covariate increased the hazard of ending a relationship, while an HR between 0.0 and 1.0 was reflected in a decreased hazard.

**Table 2. Summary of regression models used in the results chapters**

	<i>Outcome Variable</i>	<i>Estimate exponentiated?</i>	<i>Estimate that is reported</i>
<i>Linear Regression</i>	Continuous, numeric	No	$\hat{\beta}$ -coefficient ( $\beta$ )
<i>Logistic Regression</i>	Dichotomous categories	Yes	Odds Ratio (OR)
<i>Negative Binomial Regression</i>	Discrete, over-dispersed, numeric	Yes	Incidence Rate Ratio (IRR)
<i>Ordinal Regression</i>	Ordered categories	Yes	Proportional Odds Ratio (POR)
<i>Cox Proportional Hazards Model</i>	Censored event, time to event	Yes	Hazard Ratio (HR)

### 2.3.2 Multilevel models

Often researchers encounter data that include repeated measurements of outcome variables over time, or data that contain multiple measurements nested within a cluster. An example of the former would be a longitudinal study that measures participant’s blood pressure at different points in time. An instance of the latter would be a classroom that has multiple measurements of test scores for each student within that classroom. In both of those examples, researchers would expect the measurements belonging to the same participant or classroom to be more alike than measurements from individuals chosen at random in the population. This implies that each unique measurement is not completely independent from all other measurements. However, the regression strategies highlighted in Section 2.3.1 assume independence (147), and if the observations are not independent this can result in incorrect SEs.

In most of the datasets we used for analyses, the participants could report more than one relationship. For each relationship, participants reported characteristics, such as their partner's age. Like the classroom example above, we expected that these relationship characteristics would be correlated within the participant-level.

Fortunately, there are classes of models that take dependence of hierarchical data (or repeated measurements) into account when estimating SEs (148). We chose to analyse our data with generalized linear mixed-effects models (GLMMs), which are a class of multilevel/hierarchical models. Like the GLMs, GLMMs also allow modelling of non-normal response variables, using the same link functions noted above (See section 2.3.1).

GLMMs are composed of “fixed-effects” and “random-effects”. A fixed-effect is a parameter in the model for which researchers are concerned with the effects that each level has on the outcome (149). These are the covariates and main exposures/treatments in the model, and they usually have predetermined and reproducible levels. Random-effects, on the other hand, are parameters in the model whose levels are sampled from a larger population, and the levels are usually represented by unique identifiers (149). With random-effects researchers are typically interested in modelling the variation among the levels, rather than the specific effects each level has on the outcome (150). In our case, the random-effects are usually modelled for participants. We were not concerned with the effect each individual participant had on the outcome, rather we cared about the variation between and within participants, with regards to their relationship characteristics.

For all the analyses presented in the results chapters, when the unit of observation was a relationship we fitted a random intercept for the participant in the model. This means that in each model, all participants had their own regression line fitted to the data, with their own y-intercept. Thus, the intercept standard deviation estimated by the models, represented the amount of variability between participants (149). The amount of scatter of observed relationship characteristics around each participant's fitted line (predicted response), is called the residual standard deviation, and it represents the within-subject variability (149). The amount of between- and within- subject variation in partner ages was of interest to us in Chapters 5 and 7.

### 2.3.3 Non-linear covariates

When using GLMs or GLMMs, a transformation of the mean of the response variable takes place, such that the models are still linear on the logit or log scale (i.e. log odds). The assumption behind these models is that when a continuous covariate increases, the mean of the response will decrease or increase linearly. However, this presupposition can create problems if any of the continuous covariates does not have a linear relationship to the response (151).

For several of the associations explored in Chapters 5, 6, and 7 it was the case that we expected the average predicted response to fluctuate, depending on the value of the continuous covariate. For example, the age of the participant was included as a covariate in models that predict the probability of being HIV positive. We know that it is *not* the case that as age increases the probability of being HIV positive also increases linearly, so that the oldest participants have the highest probability of being HIV positive. The probability of a participant being HIV positive may increase until a certain age (e.g. 45 years old), and then subsequently decrease after that. This, in itself, may be

a simplification. In reality, it may increase, decrease, and then increase again, and we may not know at precisely which ages this happens.

There are a few ways to formulate a model so that the non-linear relationship between a continuous covariate and the response variable is made explicit. The first way to do this is to conduct a polynomial regression by adding extra covariate terms which are calculated by raising the covariate (e.g. age of participant) to a power. For example, cubic regression would use three variables for the age covariate: Age + Age<sup>2</sup> + Age<sup>3</sup>. Low-order polynomial regression rarely fits the data well, and higher-order regression has a tendency to over-fit data and generate too much variance at extreme values (152). Another common strategy in epidemiological research is to implement the continuous covariate as a step function. To do this, the continuous covariate would be split-up into discrete categories and dummy variables are produced, which are then added as predictors in the model. The downside to this method is that researchers may not know which categories make most sense, and they risk missing natural breakpoints where changes in the relationship occur (153–155).

An improvement upon the previous two methods is to use regression splines, which are more flexible than polynomials and step functions (156). To implement a regression spline for a continuous covariate, an analyst divides the range of the covariate into distinct regions. For each region, a low-degree polynomial function is fit to the observations, and the polynomials are forced to join smoothly at the boundaries between the regions – called knots. When implementing splines, researchers have to choose the number and placement of knots. Typically, this is accomplished by specifying how many knots there should be and then letting the statistical software automatically pick the placement of the knots based upon the quantiles of data points. From there  $\beta$ -coefficients for the splines are estimated.

Generalised additive models (GAMs), allow analysts to implement smoothing splines for multiple predictors (157–159). A smoothing spline is a regression spline that attempts to maximize the goodness of fit, while still maintaining parsimony, by penalizing curves that are excessively “wiggly”. In other words, GAMs attempt to achieve an optimal balance between over- and under-fitting a model. GAMs do not impose a specific covariate-response relationship, and allows the researchers to fit models to non-normal response variables, just as in GLM and GLMMs. The algorithms used in the estimation of parameters for GAMs are quite complex and this is currently an area of ongoing research (158).

We used GAMs for age and age difference covariates in Chapters 5 and 7. In Chapter 6 we used regression splines for those covariates, because GAMs had not been implemented in statistical software for CPHMs with cluster errors or mixed effects models with ordinal outcome variables. The estimates that are produced for regression and smoothing splines are often un-interpretable on their own. Therefore, analysts typically plot the smoothed curve showing the relationship between the covariate on the x-axis and the response variable on the y-axis.

### 2.3.4 Multivariate Imputation by Chained Equations

Missing data in epidemiological studies is a common problem. In sexual behaviour surveys, this typically manifests when participants refrain from answering certain questions about sexual behaviours. The customary approach of statisticians and

researchers is to delete observations that have the missing data, and use only complete cases (CC) in their analyses. Under some scenarios this practice can lead to biased estimates and it will always lead to a loss of statistical power to detect true effects.

There are typically three different processes that determine the likelihood that data will be missing, and depending on which mechanism researchers encounter, there are different strategies that can be used to minimize the bias and loss of statistical efficiency (160). The first mechanism is referred to as Missing Completely at Random (MCAR), and this happens when the missing data are not dependent on any observed or unobserved data (160). In other words, the probability that any given observation is missing, is the same for all cases. A typical example would be clerical errors related to data entry that happen randomly and result in missing observations. When this happens, a CC analysis can be done without producing any bias in the estimate of interest. However, there will still be a loss of statistical efficiency (i.e. larger confidence intervals), since the sample size will be reduced from throwing out missing observations.

The second missingness mechanism is called Missing at Random (MAR), and it occurs when observations have a high probability of being missing because they have particular values of some other observed characteristics (160). For example, people reporting a high number of lifetime sexual partners, may also be less likely to consent to HIV testing, and thus have missing HIV status in the dataset. If researchers were to conduct a CC analysis on that data, it would result in selection bias because people who have low lifetime number of partners would be over-represented in the analysis.

The final missingness mechanism is Missing Not at Random (MNAR). This results when observations on one variable have a high/low probability of being missing because of unobserved data (160). Looking at the previous example in a different way, if a participant had previously had an HIV test before entering the study and they knew their HIV status to be positive, they might be less likely to consent to HIV testing in the study (perhaps, because of the stigma associated with their positive status), and thus have a missing HIV status. This missingness pattern would appear similar to the MAR scenario. Unfortunately, there is no statistical test or way to tell the difference between an MAR and MNAR scenario, except to use prior knowledge on which mechanism seems the most likely.

When dealing with missing data that result from MAR/MNAR processes, there are several ways to impute – fill in – missing values that have been proposed over the years. These include: mean imputation, regression imputation, last observation carried forward, baseline observation carried forward, or the indicator method, to name a few. However, each of these methods are simplistic and may still result in biased estimates and SEs (161). Multiple imputation (MI), on the other hand, has been proposed to estimate plausible values for missing data using the distribution of observed data, and then adding in random noise to reflect uncertainty in the imputations and is generally valid under MAR (162).

Using MI methods to handle missing data involves three steps. The first is to produce ( $m$ ) versions of the same dataset, by simulating values for missing observations. The values generated are based upon an imputation model which regresses the variable that is missing on observed variables, and then selects values based on the posterior predictive distribution. Second, once  $m$  complete datasets are produced, they are each analysed in the same way, to produce  $m$  different estimates and variance-covariance

matrices. Third, these estimates are combined to produce one overall estimate and SE using Rubin's rules (163). The pooled variance-covariance matrix takes into account within- and between- imputation variability.

In Chapters 5 and 7, I used a technique called Multivariate Imputation by Chained Equations (MICE), to impute missing values for my data that was assumed to be MAR. This technique is useful for imputing datasets when more than one variable has missingness (161,164). MICE builds upon the MI steps outlined above, but instead of constructing one imputation model, imputation models are constructed for each variable with missing values. Through a series of iterations, a single dataset will be imputed. In the first iteration, missing values are imputed by randomly sampling one of the observed values for that variable. Then the first variable with missing values is regressed upon all the other variables in the imputation model, using only CCs. As noted above, the missing values are then filled in from the posterior predictive distribution. Then in the same iteration, the next variable with missing values is imputed in the same way. This occurs for each variable with missingness. The whole process is repeated over several iterations, to produce one imputed dataset. Ultimately,  $m$  imputed datasets will be generated in this manner (161).

MICE has the advantage of producing unbiased results with SEs that incorporate the uncertainty of imputations (162). Moreover, MICE allows imputing multiple missing variables at once, including variables of different types (e.g. continuous, or categorical) (162). However, specifying an imputation model is not trivial, and the estimates may still be biased if the imputation model is incorrect. Moreover, using standard MICE procedures is sub-optimal if researchers suspect there are non-linear associations and interactions between different variables in the analytic models. As noted in Section 2.3.3, we hypothesized that the relationships between age, age differences and HIV status in our analyses were, in fact, non-linear. Thus, we coupled a Random Forest (RF) algorithm with MICE to impute our missing data. RF is a machine learning technique that is related to and extends upon classification and regression trees (165). MICE-RF is a relatively new area of MI research, and the details of the algorithm will not be explained here. The main advantage of using RF algorithms is that they make few assumptions about the data, and have been shown to produce less biased results (166–168).

As a general rule, when I conducted MICE, I formulated the imputation model by including a combination of different socio-demographic variables (e.g. highest education level, job, etc.) and relationship characteristics (e.g. partner type, partner's place of residence, etc.) that may have been predictive of the missing values. Additionally, the imputation models included all of the variables used in subsequent analytical models. That is, they contained the exposures (e.g. age), outcomes (e.g. age of partner, or HIV status), effect modifiers (e.g. gender), and confounding covariates (e.g. race). For each analysis many imputed datasets ( $m = 50$  in Chapter 5,  $m = 100$  in Chapter 7) were produced in order to minimize uncertainty due to the randomness of imputations (169).

### 2.3.5 Thematic analysis

We used thematic analysis to explore and interpret the results from the CTSBS presented in Chapter 4. Thematic analysis is an approach for taking qualitative information and abstracting meanings and concepts from the data (170). According to Javadi and Zarea, in the context of this methodological approach a theme "...shows a

pattern or meaning related to datasets... [it is] a kind of agreement that, in comparison to the main text from which the theme is extracted, is more concise, accurate, simpler and shorter” (147, p. 34).

A thematic analysis is typically completed in six stages (172). The first stage involves the researcher familiarizing his or herself with the data by reading the transcripts or field notes while paying attention potential themes. Next, initial codes are created for specific research questions. Codes are not equivalent to themes, but rather they are just labels that refer to particular observations of interest (173). In the third phase, these codes are combined into comprehensive themes. During this stage, the researcher should define what the themes mean. Fourth, the researcher investigates how the themes tell a story about the data. In the fifth stage, the researcher further defines the themes and determines what is interesting about them. Finally, when the results are written up in a report, the researcher decides which of the themes make important new contributions and then they subsequently interpret the findings within the broader context of what they are studying.

When conducting a thematic analysis, it is not the overall prevalence of themes in the data that determines the importance of a particular theme, but rather the researcher’s assessment of what is significant (171). In reporting the findings, a researcher may decide to focus on one or a few themes that provided particularly rich details (172). This was the approach I took with the manuscript presented in Chapter 4. While the interviews encompassed a wide variety of topics about sexual relationships (See Section 2.2.4.2 above), I honed in on the themes of “age-disparate relationships” and “starting and breaking up relationships”, because they inspired stimulating discussions and original insights. Within those broad themes, I also did inductive coding. Inductive analysis means that the research derives themes from ideas emerging within the text, rather than basing them on theory (174).

Thematic analysis has several limitations that should be taken into consideration when reviewing my findings. This type of analysis is not reproducible and it is difficult to determine if the themes and interpretations of the researcher are actually compatible with the data (171). Additionally, when thematic question guides are used for interviewing, these questions can sometimes be used for coding, and thus the themes may just be based upon the researcher’s assumptions and not actual data analysis (171). Finally, the results are not generalizable and can merely be considered anecdotes, at best. The primary advantages of thematic analysis are that they are relatively easy and inexpensive to conduct and they can provide quick and easy insight into complex problems (171). An additional advantage, as I discovered in the course of my research, is that thematic contents analysis may generate new hypotheses that would not have been thought of had the researcher not conducted the analysis.

## 2.4 ETHICS APPROVAL AND CONSIDERATIONS

---

Ethics approval to collect data from participants was obtained for each of the individual studies outlined above by the universities where the principle investigators were based. For the CAPS Wave 4 data, approval was supplied by the Human Subject Review Board at Princeton University. The LNS was granted approval by the Institutional Review Boards at the University of Malawi College of Medicine, the University of Pennsylvania, and Columbia University Medical Centre. Ethics committees at the National Health

Sciences Research Council (Malawi), the University of California at San Diego, and George Washington University approved the SIHR study. We applied for and received ethics approval for the CTSBS sexual behaviour survey and qualitative interviews from the Stellenbosch University Health Research Ethics Committee. Finally, the same committee also granted us separate approval to conduct secondary analyses of age-mixing data.

Since we conducted the data collection for the CTSBS study, I will briefly highlight some of our ethical considerations. All participants were provided with information about the study and were free to ask questions before they supplied their written informed consent. We also explained to participants that if they refused to participate there would be no negative consequences, and they could withdraw at any point in time. Likewise, we did not provide any financial or material incentives for participating. Participants were guaranteed complete anonymity and we explained that personal information would be kept confidential. I, personally, trained all of the research assistant fieldworkers and I encouraged them to make the participants feel comfortable and safe. No biomarker data were collected, so this was considered to be a low-risk study.

In addition to the previously mentioned ethical considerations for the CTSBS, the qualitative component posed a few more risks. Discussing dynamics of relationships and other sexual information may be a source of discomfort for participants. I tried to be cognizant of any signs of distress. Fortunately, all of the participants in the interviews seemed willing and psychologically sound before, during, and after completion of the interviews. Additionally, as a researcher conducting research in a disadvantaged area with a high prevalence of violent crime, I was also aware of safety concerns posed to myself and research assistant. Throughout the duration of the study, I made consistent efforts to locate safe, private and enclosed spaces to conduct interviews. Moreover, I did not walk in any locations without a local person to guide me to safe locations.

For the purposes of our secondary statistical analyses our primary ethical considerations were related to maintaining the security of the datasets that we used. We kept all copies of the datasets on password-protected computers or private online repositories. The datasets we were given – and including the CTSBS study we conducted – were all anonymized with uniquely generated ID variables for each person. We did not have access to any identifying information (e.g. ID numbers, names, GPS coordinates, and addresses) during the analyses, so we could not harm participants by linking exposures or outcomes to specific individuals. There were no immediate benefits or risks to the participants who were included in our secondary analyses because we did not make any contact with them.

## 2.5 DISSEMINATION OF RESULTS

---

The foundation for this dissertation is comprised of the following manuscripts, published or submitted to international peer-reviewed journals:

- 1) **Beauchair, R.;** Kassanjee, R.; Welte, A.; Temmerman, M.; Delva, W. (2012). 'Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa.' *The European Journal of Contraception and Reproductive Health Care* 17(1):30.

- 2) **Beauchair, R.** and Delva, W. (2013) 'Is younger really safer? A qualitative study of perceived risks and benefits of age-disparate relationships among women in Cape Town, South Africa'. *PloS ONE* 8(11): e81748.
- 3) **Beauchair, R.;** Helleringer, S.; Hens, N.; Delva, W. (2016) 'Age differences between sexual partners, behavioural and demographic correlates, and HIV infection on Likoma Island, Malawi'. *Scientific Reports*: 6.
- 4) **Beauchair, R.;** Dushoff, J.; Delva, W. (2018) 'Partner age differences and associated sexual risk behaviours among adolescent girls and young women in a cash transfer programme for schooling in Malawi'. *BMC Public Health*: Under Review.
- 5) **Beauchair, R.;** Hens, N.; Delva, W. (2018) 'Age-mixing among HIV positive and negative men and women: population and individual-level comparisons from a Cape Town sexual behaviour survey'. *Epidemics*: Under Review.

Some of the results have also been presented at international conferences:

- 1) **Beauchair, R.;** Helleringer, S.; Hens, N.; Delva, W. 'Age-mixing patterns and HIV transmission in the Likoma Network Study'. *AIDS 2016*, 18-22 July 2016: Durban, South Africa. (Poster)
- 2) **Beauchair, R.** and Delva, W. 'The Lessor of Two Evils: A Qualitative Study of Perceived Risks of Age-disparate Relationships Among Women in Cape Town, South Africa'. *6<sup>th</sup> SA AIDS Conference*, 18-21 June 2013: Durban, South Africa. (Oral)
- 3) **Beauchair, R.;** Dushoff, J.; Delva, W. 'Partner age differences and associated sexual risk behaviours among adolescent girls and young women in a cash transfer programme for schooling in Malawi'. *Epidemics*<sup>6</sup>, 29 Nov-01 December 2017: Sitges, Spain. (Poster)
- 4) **Beauchair, R.;** Hens, N.; Delva, W. 'Age-mixing among HIV positive and negative men and women: population and individual-level comparisons from a Cape Town sexual behaviour survey'. *Epidemics*<sup>6</sup>, 29 Nov-01 December 2017: Sitges, Spain. (Poster)

Finally, some of the results from the CTSBS qualitative study were disseminated to the study communities in a free, public event held on 27 April 2015 in Khayelitsha, Cape Town<sup>1</sup>. Members of the community could come, eat a snack, enjoy music played by a local DJ, peruse pamphlets, and listen to speakers, including the principle investigator, Prof. Wim Delva and Shiela McCloen, Deputy Director of comprehensive health programmes in Khayelitsha. This initiative was part of our efforts to be more inclusive of the study communities in the distribution of scientific knowledge. We also encouraged open dialogue with community members and had a discussion with them about their interpretation of the results. Please see Appendix E for an English translation of the brochure that was distributed.

---

<sup>1</sup> Other study results were also shared with the public, many of which are not featured in this dissertation. The manuscripts, from which those results come, are in Appendices A-D.

CHAPTER 3: AGE-DISPARATE RELATIONSHIPS AND  
IMPLICATIONS FOR STI TRANSMISSION AMONG  
YOUNG ADULTS IN CAPE TOWN, SOUTH AFRICA

---



Photo credit: Paul Saad via Visual Hunt / CC BY-NC-ND

### 3.1 AIMS OF ANALYSIS

---

The manuscript that follows presents the first study I conducted that attempted to understand age differences in relationships. In 2012, when the manuscript was published in *The European Journal of Contraception and Reproductive Health Care*, little was known about age-mixing patterns in real-world sub-Saharan African populations. We, therefore, primarily focused on calculating descriptive statistics and applying naïve regression modelling strategies to generate further hypotheses about the causal pathways between ADR and HIV acquisition.

We used the CAPS data to: 1) Estimate the prevalence of ADR and visualize the age-mixing patterns among black and coloured participants in this representative sample of young adults from Cape Town, South Africa (**Overall Objective 1**, See Section 1.6); 2) Identify socio-demographic predictors of participants who engaged in ADR (**Overall Objective 2**); and 3) Compare individual and relationship characteristics between women who were in ADR versus those who were not in ADR (**Overall Objectives 2 and 3**).

## Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa

Roxanne Beauclair\*, Reshma Kassarjee\*, Marleen Temmerman†, Alex Welte\* and Wim Delva\*†

\*South African Centre for Epidemiological Modelling & Analysis, Stellenbosch University, South Africa, and

†International Centre for Reproductive Health, Ghent University, Ghent, Belgium

**ABSTRACT** **Objectives** To estimate the prevalence of age-disparate (AD) relationships among young black and coloured adults in Cape Town (South Africa) and determine socio-demographic predictors and individual and relationship characteristics of women in these relationships. **Methods** A secondary analysis of the Cape Area Panel Study ( $N=1960$ ) data was conducted. Descriptive statistics were used to quantify the age-mixing pattern and logistic regression was used to identify significant socio-demographic and behavioural correlates of AD relationships. **Results** Prevalence of AD relationships was high in both black (36%) and coloured (28%) women. The average age difference between male respondents and their partners increased with age. Young, black women who spent fewer nights under the same roof in one week, had a deceased parent, and were not currently attending classes were more likely to be in an AD relationship. Reports of sexually-transmitted infection (STI) symptoms in the last month and unprotected sex were more common among women in AD relationships. **Conclusions** AD relationships are common among young women in Cape Town. Home and family stability is preventative of young women engaging in AD relationships. Therefore, holistic, societal interventions may reduce AD relationships, which are a risk factor for STIs.

**KEY WORDS** Age disparity; Inter-generational sex; HIV prevention; South Africa; Sexual behaviour; Sexually transmitted infection

### INTRODUCTION

Recent evidence suggests that engaging in age-disparate (AD) relationships – a relationship where the male partner is five or more years older than the female partner – is a strong risk factor for HIV infection among girls

and young women in sub-Saharan Africa<sup>1–4</sup>. In the 2005 South African national HIV prevalence survey, the HIV prevalence among girls aged 15–19 with partners five or more years older was 30%. The prevalence among girls from the same age group in

Correspondence: Roxanne Beauclair, SACEMA, c/o StIAS, Private Bag X1, Matieland, Stellenbosch, Republic of South Africa, 7602. Tel: +27 (0)21 808 2589. Fax: +27 (0)21 808 2586. E-mail: roxanne.beauclair@gmail.com

relationships where the age difference was less than five years was 17%. For women aged 20–24, a similar discrepancy in HIV prevalence was observed (35% versus 23%)<sup>5</sup>.

For young women in South Africa, older partners are more likely to expose them to the virus, since the HIV prevalence in men increases with age, peaking above 25% among men aged 30–34<sup>6</sup>. Moreover, levels of condom use may be lower in older men who have sexual relationships with younger women than in men of the same age who do not engage in AD relationships<sup>7,8</sup>.

In a study conducted in Harare, Zimbabwe, Wyrod *et al.* found that the largest age gaps in relationships occurred between older men and younger female partners of steady, long-term relationships and condom use within these relationships was rare<sup>9</sup>. Further complicating this common Southern African scenario, older male partners of young women may be more likely to partake in multiple, concurrent relationships with casual partners while maintaining 'long-term' partnerships<sup>10</sup>.

The importance of age disparity for the acquisition of HIV and other sexually transmitted infections (STIs) in young women critically depends on the prevalence of AD relationships, as well as the current and past sexual risk behaviour of men and women who engage in such relationships. In this analysis, we explore the age-mixing pattern and calculate the prevalence of age disparity in young, unmarried people in Cape Town, South Africa. Further, we identify socio-demographic predictors of young women being in AD relationships. Lastly, we compare individual and relationship characteristics between young women who are engaged in AD relationships with those who are not.

## METHODS

This study uses data obtained from the fourth wave of the Cape Area Panel Study (CAPS), a cohort study of young adults in metropolitan Cape Town that explored various life transitions by administering questionnaires about reproductive and sexual health, employment, and family dynamics. Wave 1 of the study began in August 2002, and approximately 4,800 adolescents and young adults, aged 14–22 years, were sampled from 5256 households. Wave 2 followed up

a third of participants from Wave 1 in 2003 and the remaining two-thirds of the youths in 2004. Wave 3 youths were re-interviewed in 2005 and again for Wave 4 in 2006. A detailed description of the study design and sampling methods is provided elsewhere<sup>11,12</sup>. This analysis only uses the responses to the Wave 4 interviews, which followed up as many of the original participants as possible, resulting in a sample size of 3,439 young adults, comprising 72% of the original sample.

Participants were excluded from the analysis if they identified themselves as white or Indian because of insufficient sample sizes for these sub-populations. Additionally, 166 observations were excluded due to inconsistent ( $n = 136$ ) or missing ( $n = 30$ ) age data. Also, we only included adolescents and young adults who had ever had sex because we wanted our sample to be composed of sexually active participants. Finally, we excluded participants who were married since we had insufficient data on the age of their spouses and their age-mixing patterns are likely to be different from those of unmarried people. This resulted in a final sample size of 1960 adolescents and young adult men and women aged 17–26. The analyses we conducted were stratified by sex and race, resulting in final sample sizes of 570 black men, 693 black women, 323 coloured men, and 374 coloured women. Race stratification (black/coloured) was conducted because it was thought to be an effect modifier as it is considered a proxy for access to various economic, social and health opportunities in South Africa. The term 'coloured' is used in South Africa to refer to racially-mixed descendants of European, Asian and indigenous African populations.

We determined whether or not the respondent was currently in an AD relationship by finding the age difference between the respondent and his or her partner. Relationships in which the female partner was five or more years younger than the male partner, were classified as age-disparate.

Nine socio-demographic variables were assessed as potential predictors of age disparity. We explored the effect of the age of the respondent (continuous, 17–26 years), total monthly earnings of the respondent (no income, 0–500 ZAR [0–45 €] per month, > 500 ZAR [ > 45 €] per month), having completed grade 12 of high school (yes/no) and obtained a degree (yes/no). Home stability was operationalised by the following socio-demographic variables: survival

of mother (deceased, alive, unknown), survival of father (deceased, alive, unknown), currently attending classes (yes/no), and the number of nights spent under their household roof in a week (count, 0–7). These three variables were chosen in order to determine whether home stability made a woman more likely to engage in AD relationships. For our final socio-demographic variable, we developed an indicator of the respondents' overall well-being (score, 0–24) based on six questions relating to the frequency of experiencing negative emotions over the preceding 30 days ('nervous', 'hopeless', 'restless or fidgety', 'depressed', 'everything was an effort', 'worthless'). A higher score was indicative of an increased sense of well-being.

For the comparison of women in and out of AD relationships we considered the presence or absence of the following behavioural characteristic variables: condom used at last sexual intercourse, STI symptoms in the last 30 days, drugs used in the last 30 days, alcohol used in the last 30 days, monetary/in-kind transfers from partner, and cohabitation with partner. We also compared the number of partners in the past 12 months (<2 partners or ≥2 partners) and number of lifetime partners (<4 partners or ≥4 partners) between both groups of women.

Statistical analyses were performed using Stata version 10 (Stata Corporation, College Station, TX). We calculated the proportions of respondents engaging in AD relationships by age category, gender and race at the time of the study, as well as, at sexual debut. Scatter plots were used to visually depict the age-mixing patterns. Average age differences and standard deviations of the partner age were calculated to quantify and compare these patterns. Multiple logistic regression models were used to evaluate potential socio-demographic predictors of women engaging in AD relationships. Finally, we calculated unadjusted odds ratios (ORs) to compare individual and relationship characteristics of women in and out of AD relationships.

Ethical approval for wave 4 of this study was granted by the human subject review boards at Princeton University, the University of Michigan, and the University of Cape Town. Written, informed consent was received from all respondents and written parental consent was obtained for all interviews with respondents under the age of 18.

## RESULTS

Of the 1960 respondents in our sample 1067 were women (black,  $n = 693$ ; coloured,  $n = 374$ ). Among all men and women, 1550 indicated that they currently had a partner (79%), and 1513 of those provided an age for their partner (98%). Of these, 39% ( $n = 583$ ) were black women and 17% ( $n = 260$ ) were coloured women. The distribution of socio-demographic and behavioural characteristics in women did not appear to be the same across the races (Table 1). Fewer black women had living mothers and fathers, made more than 500 ZAR (>45 €) per month, used drugs, used alcohol, and cohabited with their partner, compared to coloured women. Coloured women had a lower prevalence of currently attending class, using a condom at last sex, receiving monetary or in-kind transfers from their partner, and had a lower average number of lifetime partners.

### Prevalence of age disparity

AD relationships were common in both black and coloured women (36% and 28%, respectively). Among women of both races, the prevalence of AD relationships varied across the age groups, but never dropped below 20%, and peaked at 40% in black women and 34% in coloured women (Figure 1A). In both black and coloured men the prevalence of AD relationships started off low and gradually increased to 37% and 39% in 25–26 year olds, respectively (Figure 1B). At sexual debut, 16% of black women and 22% of coloured women were engaged in AD relationships. In contrast, only 1% of black and 0% of coloured men engaged in AD relationships at sexual debut.

### Age-mixing patterns

The average age of the partners of 18-year-old black men was 17 and the average age difference increased by 0.34 years for every one year age increase, so that the average partner age was 22 for 26-year-old black men. For coloured men, the same average age difference was witnessed at age 18, but this average difference only started to increase in relationships with men older than 23 years (Figure 2A and 2B). The age-mixing patterns in black and coloured women were similar: both groups of women had an

**Table 1** Age disparity, socio-demographic predictors, and relationship characteristics of adolescent and young adult women, by race.

Variable	Adolescent and young adult women		p-values for difference by race**
	Black	Coloured	
Number of respondents (Total = 1067)	693	374	
Mean age of respondent (SD)	21.74 (2.42)	21.76 (2.36)	0.89
Age-disparity at sexual debut (%)	15.49	21.59	0.03*
Age-disparity in current relationship (%)	36.02	27.69	0.02*
Mean nights under roof in week (SD)	6.69 (1.08)	6.93 (0.58)	<0.01*
Mean well-being score (SD)	20.97 (4.57)	21.79 (4.41)	0.02*
Survival of mother (%)			0.01*
Deceased	12.46	6.25	
Alive	87.02	92.97	
Unknown	0.53	0.78	
Survival of father (%)			<0.01*
Deceased	27.89	20.31	
Alive	61.75	75.78	
Unknown	10.35	3.91	
Total monthly earnings (%)			<0.01*
No income	26.14	18.75	
0–500 ZAR (0–45 €) per month	38.07	16.41	
> 500 ZAR (>45 €) per month	35.79	64.84	
Completed grade 12 of high school (%)	37.02	42.97	0.11
Obtained high school diploma or college degree (%)	10.18	8.59	0.52
Currently attending classes (%)			<0.01*
Not attending	59.65	89.06	
Attending	29.12	6.64	
Unknown	11.23	4.30	
Condom use at last sexual intercourse (%)	63.12	29.41	<0.01*
STI symptoms in last 30 days (%)	4.12	1.92	0.15
Alcohol use in last 30 days (%)	14.63	44.23	<0.01*
Drug use in last 30 days (%)	0.69	3.46	<0.01*
Monetary/in-kind transfer (%)	56.35	44.98	<0.01*
Cohabitate with partner (%)	6.69	11.15	0.04*
Mean number of partners in the last 12 months (SD)	1.20 (0.61)	1.07 (0.69)	0.01*
Mean number of partners in lifetime (SD)	3.10 (2.05)	1.76 (1.53)	<0.01*

\*Indicates a significant difference between races at a 5% significance level.

\*\*Calculated using Fisher's exact test.

SD, standard deviation.

average partner age around 21 years when they were 18 and 30 years when they were 26 (Figure 2C and 2D). For men of both races, the variance of partner age was small in the youngest men, and increased with age. However, for women, the variance of partner age was large at all ages (Table 2).

### Socio-demographic predictors of AD relationships in women

Table 3 shows the adjusted odds ratios (ORs), based on logistic regression models, for potential socio-demographic predictors of AD relationships for black

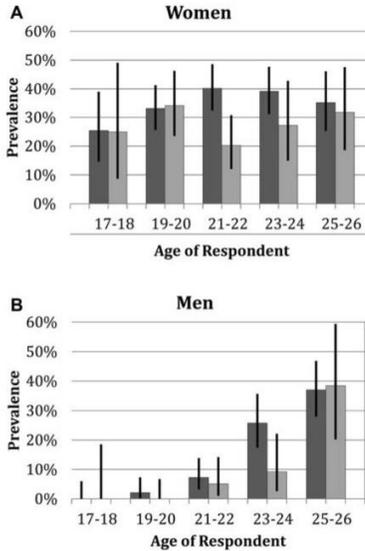


Figure 1. Prevalence of age-disparate relationships among different age categories of (A) women, and (B) men. The dark grey bars show the prevalence in black respondents, while the light grey ones represent coloured respondents. The lines show the 95% confidence intervals for the prevalences.

and coloured women. For black women, the increased number of nights spent in her household during the week, survival of the mother, survival of the father, and attending classes were all significant and protective against being in an AD relationship. For coloured women, only the well-being score was found to be protective against being in an AD relationship.

#### Comparison of individual and relationship characteristics

Table 4 illustrates the differences in individual and relationship characteristics for those in an AD relationship compared to those who were not, and ORs that were calculated for each of these characteristics. Black women in an AD relationship were more likely to report STI symptoms in the past 30 days: 7% had STI symptoms compared to 2% of those not in

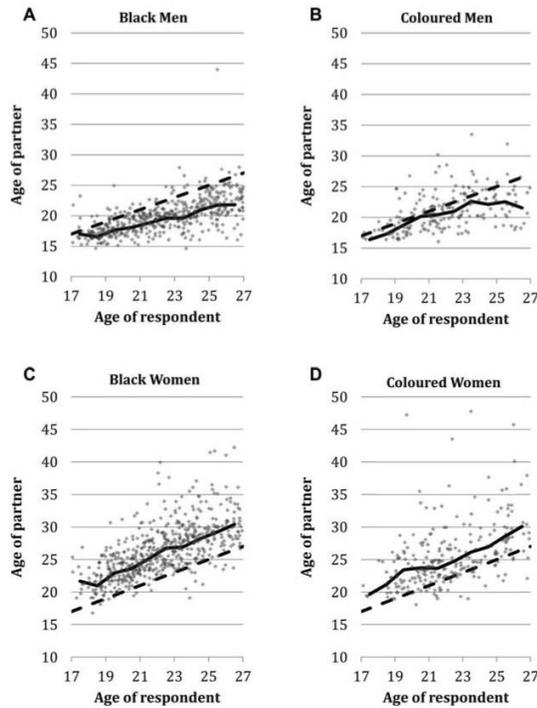
AD relationships. Being in an AD relationship also increased the likelihood of receiving monetary support or in-kind transfers from the current partner: 64% compared to 52%. Of those in an AD relationship, 11% were cohabiting with their partners compared to 5% of those not in an AD relationship. Finally, the proportion of black women with fewer than four lifetime partners was 64% among those in an AD relationships, compared to 73% of those not in an AD relationship.

Amongst coloured women, those in an AD relationship were less likely to use condoms at the last sexual intercourse: of those in AD relationships, only 19% used a condom compared to 33% of those not in an AD relationship. As in black women, STI symptoms in the past 30 days were more prevalent in AD relationships.

#### DISCUSSION

Our analysis of the CAPS Wave 4 responses scrutinises the age-mixing patterns and characteristics of AD relationships among adolescents and young adults in Cape Town, South Africa. Adolescent and young adult men have increasing access to younger female partners and greater partner age variance as they get older. Additionally, our analysis provides direct and indirect evidence for age disparity as a risk factor for STI in young women, and suggests that a stable home environment, for young black women, is protective against engaging in AD relationships.

Given the important historic and present differences in socio-demographic characteristics of young black and coloured people in South Africa, the key limitation of our analysis is that the race-stratified analysis was under-powered for coloured respondents, due to a small sample size. Hence, our interpretation of the results focuses on identifying empirical evidence to support general hypotheses about causal links between AD relationships and the risk of STIs, including HIV, rather than on differences in the factors associated with AD relationships between the two races. Perhaps the greatest strength of our study is that it provides a detailed, description of the age-mixing patterns, the prevalence of AD relationships, along with socio-demographic predictors of age disparity and characteristics of AD relationships relevant for STI transmission, all based on



**Figure 2.** Age of the current partner as a function of the age of the respondent for (A) black men, (B) coloured men, (C) black women, and (D) coloured women. The solid line indicates the average age of the partner when the respondent is aged a given number of years. The dashed line represents no age difference between the respondent and his or her partner.

the same data source. Previous studies have focused primarily on either predictors, prevalence or consequences of age-disparity<sup>1,8,9,13,14</sup>, have collated evidence from diverse settings and study populations<sup>3</sup>, or have provided mainly qualitative insights<sup>2,15,16</sup>.

We confirm previous findings that as young adult men get older and more sexually experienced the average age gap between them and their partners increases<sup>13</sup>. Consequently, AD relationships are uncommon in adolescent men, but become more frequent for men in their twenties. In contrast, the average age gap for relationships of young women is large at sexual debut and continues to stay large as they become older.

Our results also demonstrate that men have partners from a wider age range as they grow older. The increasing average age difference between men and their partners and the increasing variance around their partners' average age have important implications for transmission of HIV and other STIs. They signify critical transmission pathways from older to younger female birth cohorts with men acting as a vector of transmission. Such pathways are vital for sustained transmission of STIs, especially if the per-sex-act probability of transmission is low, such as for HIV.

Additionally, our study shows that a greater proportion of those in an AD relationship had STI symptoms

Table 2 Average partner age and standard deviation as a function of the age of respondent.

Age of respondent	Women		Men	
	Black Average partner age (SD) (n = 583)	Coloured Average partner age (SD) (n = 260)	Black Average partner age (SD) (n = 471)	Coloured Average partner age (SD) (n = 199)
18	20.95 (2.17)	21.18 (2.40)	16.56 (1.06)	17.31 (0.95)
19	22.93 (2.45)	23.41 (5.27)	17.65 (1.06)	17.31 (0.95)
20	23.59 (2.47)	23.73 (3.53)	18.10 (1.57)	20.14 (1.93)
21	25.10 (2.42)	23.65 (3.16)	18.82 (1.49)	20.42 (2.99)
22	26.72 (3.49)	24.76 (4.68)	19.59 (1.85)	21.00 (2.80)
23	26.91 (3.04)	26.16 (5.55)	19.60 (2.40)	22.60 (3.25)
24	28.20 (3.12)	26.95 (3.85)	20.95 (2.25)	22.11 (2.08)
25	29.25 (4.05)	28.61 (4.65)	21.76 (3.40)	22.53 (3.50)
26	30.41 (3.56)	30.09 (5.47)	21.82 (2.72)	21.55 (3.05)

SD, standard deviation.

than those who were not, thereby strengthening the evidence base for AD as a risk factor for STI acquisition in young women<sup>13,14</sup>. The risk of STI acquisition is largely determined by the probability of STI infection in the partner, the frequency of sex acts and the level of condom use in a relationship<sup>17,18</sup>. Black and coloured women in AD relationships were respectively three and eleven times more likely to report STI symptoms in the last 30 days than their peers in non-AD relationships. Since the prevalence of most STIs in men increases with age, at least until their early thirties<sup>6</sup>, the probability of the male partner being infected with one or more STIs is increased in AD relationships<sup>13</sup>. Moreover, cumulative STI risk exposure due to multiple unprotected sex acts may be higher in AD relationships. No condom use at last sex was more frequently reported by women in AD relationships, as was cohabitation with the partner, a factor known to be associated with higher sex frequency and lower condom use<sup>19,20</sup>.

For black women, spending more nights in the same household in a given week, having both mother and father alive, and attending classes were all preventive factors for engaging in an AD relationship. These factors seem to point to familial and home stability as a key deterrent when choosing to participate in an AD relationship. Among coloured women, those who had a higher well-being score had a tendency to stay away from AD relationships. We hypothesise that a higher emotional well-being, as well as increased family and

home stability, allows young women to pursue and be more comfortable in relationships with younger male partners closer to their own age because they do not need to rely on them as much for financial or emotional support. This would also explain why we found that black women in AD relationships are more likely to be in relationships where they have accepted monetary or in-kind support. Recent studies have pointed out that women enter AD relationships for a variety of reasons. Often it is for access to material goods<sup>15,21</sup> or help when there is food insufficiency at home<sup>22</sup>. Women have also been known to seek these relationships purely for the opportunity to feel loved and wanted, which helps to affirm their sense of self-worth and boost their self-esteem<sup>2</sup>. We support Leclerc-Madlala's view that the emotional and financial benefits women receive from engaging in AD relationships may outweigh the perceived risks of STI acquisition<sup>23</sup>.

Too many HIV and STI prevention programmes for young people in sub-Saharan Africa adopt narrow approaches to sexual and reproductive health that only address behaviour change, ignoring the underlying psycho-social roots of relationship decisions. Unsurprisingly, these programmes have had little or no impact on sexual risk behaviour<sup>24</sup>. Our study is indicative of the need for operational research that assesses the effect of more holistic, structural interventions on engaging in AD relationships and the subsequent risks of STIs.

**Table 3** Adjusted odds ratios for socio-demographic characteristics by race in women.

	Blacks (n = 583) Adjusted Odds Ratios (ORs)*		Coloured (n = 260) Adjusted Odds Ratios (ORs)*	
	OR	95% CI	OR	95% CI
Age	0.96	0.88–1.06	1.07	0.93–1.23
Nights under roof in week	0.80**	0.68–0.94	1.14	0.63–2.06
Well-being score	0.99	0.95–1.03	0.93**	0.88–0.99
Mother alive				
Deceased	1.00		1.00	
Alive	0.59**	0.35–1.00	1.80	0.47–6.85
Unknown			5.25	0.19–144.82
Father alive				
Deceased	1.00		1.00	
Alive	0.58**	0.39–0.87	1.94	0.86–4.39
Unknown	0.96	0.51–1.78	0.99	0.16–6.02
Total monthly earnings				
No income	1.00		1.00	
0–500 ZAR (0–45 €) per month	1.15	0.72–1.84	1.71	0.67–4.36
> 500 ZAR (> 45 €) per month	1.39	0.85–2.28	0.79	0.37–1.72
Completed grade 12 of high school				
No	1.00		1.00	
Yes	0.91	0.60–1.36	0.78	0.40–1.53
Obtained degree/diploma				
No	1.00		1.00	
Yes	0.86	0.45–1.65	0.92	0.29–2.94
Currently attending classes				
No	1.00		1.00	
Yes	0.58**	0.36–0.94	0.39	0.08–1.81
Unknown	0.66	0.36–1.21	1.79	0.48–6.71

\*Adjusted for all variables shown in the table.

\*\*Indicates an odds ratio significantly different from 1.00 at a 5% significance level.  
CI, confidence interval.

Parallel with operational HIV prevention research, more quantitative analyses of age-mixing patterns in settings with a high HIV prevalence are needed to examine the complex correlations between age disparity and other facilitators of STI transmission: most importantly the frequency of partner change, concurrency, oscillating migration, condom use, and patterns of alcohol use. To this end, a cross-sectional survey in three urban disadvantaged communities in the greater Cape Town area is currently exploring respondents' sexual and socio-demographic histories<sup>25</sup>. Finally, a computer simulation study is underway to quantify the facilitating effect that the age-mixing patterns described above may have on the transmission dynamics of HIV

infection and other STIs among young adults in Cape Town.

#### ACKNOWLEDGEMENTS

This analysis was conducted with financial support from the European Society of Contraception and Reproductive Health (ESC), the Flemish Research Fund (FWO), and the Canadian International Development Agency (CIDA).

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Eur J Contracept Reprod Health Care. Downloaded from informahealthcare.com by 196.46.71.251 on 01/12/12  
For personal use only.

**Table 4** Comparison of individual and relationship characteristics for adolescent and young adult women in AD relationships to those who are not, by race.

	Black (n = 583)				Coloured (n = 260)			
	No AD relationship (%)	AD relationship (%)	OR (95%CI)	p-value**	No AD relationship (%)	AD relationship (%)	OR (95%CI)	p-value**
Condom use at last sex	64.61	60.48	0.84 (0.58–1.21)	0.33	33.33	19.44	0.48 (0.23–0.96)	0.03*
STI symptoms	2.41	7.14	3.11 (1.25–8.20)	0.01*	0.53	5.56	11.00 (1.05–544.44)	0.02*
Alcohol use	13.71	16.27	1.22 (0.74–2.01)	0.46	46.28	38.89	0.74 (0.41–1.33)	0.33
Drug use	0.54	0.95	1.78 (0.13–24.75)	0.62	3.72	2.78	0.74 (0.07–4.01)	1.00
Monetary support or in-kind transfers	52.39	63.83	1.60 (1.10–2.35)	0.01*	45.29	44.07	0.95 (0.50–1.80)	0.88
Cohabitate	4.56	10.48	2.45 (1.21–5.04)	0.01*	9.04	16.67	2.01 (0.82–4.76)	0.12
Number of partners in last 12 months (<2 partners)	85.41	82.30	0.79 (0.49–1.29)	0.34	93.79	88.24	0.49 (0.17–1.49)	0.18
Number of partners in lifetime (<4 partners)	71.69	63.20	0.69 (0.47–1.00)	0.05*	93.37	88.60	0.57 (0.19–1.83)	0.28

\*Indicates a significant difference between relationships with and without AD at a 5% significance level.

\*\*Calculated using Fisher's exact test.

AD, age disparate; OR, odds ratio; CI, confidence interval; STI, sexually transmitted infection.

## REFERENCES

- Kelly RJ, Gray RH, Sewankambo NK, et al. Age differences in sexual partners and risk of HIV-1 infection in rural Uganda. *J Acquir Immune Defic Syndr* 2003;32:446–51.
- Leclerc-Madlala S. Age-disparate and intergenerational sex in southern Africa: The dynamics of hypervulnerability. *AIDS* 2008;22(Suppl. 4):S17–25.
- Luke N. Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa. *Stud Fam Plann* 2003;34:67–86.
- Sa Z, Larsen U. Gender inequality increases women's risk of HIV infection in Moshi, Tanzania. *J Biosoc Sci* 2008;40:505–25.
- Shisana O, Rehle T, Simbayi LC, et al. *South African national HIV prevalence, HIV incidence, behaviour and communication survey, 2005*. Cape Town: HSRC Press 2005.
- Shisana O, Rehle T, Simbayi L, et al. *South African national HIV prevalence, incidence, behaviour and communication survey 2008: A turning tide among teenagers?* Cape Town: Human Sciences Research Council 2009.
- Bankole A, Ahmed FH, Neema S, et al. Knowledge of correct condom use and consistency of use among adolescents in four countries in sub-Saharan Africa. *Afr J Reprod Health* 2007;11:197–220.
- Luke N. Confronting the 'sugar daddy' stereotype: Age and economic asymmetries and risky sexual behavior in urban Kenya. *Int Fam Plann Perspect* 2005;31:6–14.
- Wyrod R, Fritz K, Woelk G, et al. Beyond sugar daddies: Intergenerational sex and AIDS in urban Zimbabwe. *AIDS Behav* 2011;15:1275–82.
- Chopra M, Townsend L, Johnston L, et al. Estimating HIV prevalence and risk behaviors among high-risk heterosexual men with multiple sex partners: Use of respondent-driven sampling. *J Acquir Immune Defic Syndr* 2009;51:72–7.
- Kenyon C, Dlamini S, Boule A, et al. A network-level explanation for the differences in HIV prevalence in South Africa's racial groups. *Afr J AIDS Res* 2009;8:243–54.
- Tekorang EY, Maticka-Tyndale E, Fajulton F. A multi-level analysis of risk perception, poverty and sexual risk-taking among young people in Cape Town, South Africa. *Health & Place* 2011;17:525–35.
- Gregson S, Nyamukapa CA, Garnett GP, et al. Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *Lancet* 2002;359:1896–903.
- Katz I, Low-Beer D. Why has HIV stabilized in South Africa, yet not declined further? Age and sexual behavior patterns among youth. *Sex Transm Dis* 2008;35:837–42.
- Cockcroft A, Kunda JL, Kgakole L, et al. Community views of inter-generational sex: Findings from focus groups in Botswana, Namibia and Swaziland. *Psychol Health Med* 2010;15:507–14.
- Nkosana J, Rosenthal D. The dynamics of intergenerational sexual relationships: The experience of schoolgirls in Botswana. *Sex Health* 2007;4:181–7.
- Garnett GP, White PJ, Ward H. Fewer partners or more condoms? Modelling the effectiveness of STI prevention interventions. *Sex Transm Infect* 2008;84 (Suppl. 2):ii4–11.
- Rottingen JA, Garnett GP. The epidemiological and control implications of HIV transmission probabilities within partnerships. *Sex Transm Dis* 2002;29:818–27.
- Dunkle KL, Stephenson R, Karita E, et al. New heterosexually transmitted HIV infections in married or cohabiting couples in urban Zambia and Rwanda: An analysis of survey and clinical data. *Lancet* 2008;371:2183–91.
- Polacek M, Celentano DD, O'Campo P, Santelli J. Correlates of condom use stage of change: Implications for intervention. *AIDS Educ Prev* 1999;11:38–52.
- Underwood C, Skinner J, Osman N, Schwandt H. Structural determinants of adolescent girls' vulnerability to HIV: Views from community members in Botswana, Malawi, and Mozambique. *Soc Sci Med* 2011;73:343–50.
- Weiser SD, Leiter K, Bangsberg DR, et al. Food insufficiency is associated with high-risk sexual behavior among women in Botswana and Swaziland. *PLoS Med* 2007;4:1589–97; discussion 1598.
- Leclerc-Madlala S. Transactional modernity sex and the pursuit of modernity. *Soc Dynamics* 2004;2:1–21.
- Michielsen K, Cherish M, Luchters S, et al. Effectiveness of HIV prevention for youth in sub-Saharan Africa: Systematic review and meta-analysis of randomized and nonrandomized trials. *AIDS* 2010;24:1193–202.
- Delva W, Beaulair R, Welte A, et al. Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: A study protocol. *BMC Public Health* 2011;2:616.

CHAPTER 4: IS YOUNGER REALLY SAFER? A  
QUALITATIVE STUDY OF PERCEIVED RISKS AND  
BENEFITS OF AGE-DISPARATE RELATIONSHIPS  
AMONG WOMEN IN CAPE TOWN, SOUTH AFRICA

---



Photo credit: David Rosen Photography via Visual hunt / CC BY

## 4.1 AIMS OF ANALYSIS

---

In the previous chapter, I noted that young women from the CAPS population who were not in ADRs tended to have higher emotional well-being scores and more family and home stability. We argued that this would have allowed the women to pursue relationships with age-similar men, who would not be expected to provide as much monetary and emotional support to the young women. Following the completion of that study and the CTSBS survey we wanted to understand whether women in disadvantaged communities would share this perspective. Thus, using the qualitative data from the CTSBS, we aimed to identify some motivations for engaging in ADR. Additionally, we wanted to explore how women perceived the risks of engaging in ADR versus age-similar relationships and how those risk perceptions may influence decisions to begin or end relationships. Jointly, these specific study objectives aided in achieving my **overall objective 2** outlined in Section 1.6 by providing a richer description of the context and circumstances that lead women to choose ADR over age-similar relationships. The following manuscript was published in *PloS ONE*.

## Is Younger Really Safer? A Qualitative Study of Perceived Risks and Benefits of Age-Disparate Relationships among Women in Cape Town, South Africa

Roxanne Beauclair<sup>1\*</sup>, Wim Delva<sup>1,2</sup>

**1** The South African Department of Science and Technology/National Research Foundation (DST/NRF) Centre of Excellence in Epidemiological Modeling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa, **2** International Centre for Reproductive Health, Ghent University, Ghent, Belgium

### Abstract

Young women in age-asymmetric relationships may be at an elevated risk for acquisition of HIV, since relationships with older men are also correlated with other risk behaviors like less condom use. Qualitative studies have shown that women are motivated to participate in these relationships for money and emotional support. However, there is a paucity of research on women's perceived risks of these relationships, particularly in South Africa. To this end, we conducted in-depth interviews with 23 women recruited from three urban communities in Cape Town. A thematic question guide was used to direct the interviews. Thematic content analysis was used to explore women's perceived risks of age-disparate and non-age-disparate relationships, the benefits of dating older men, and risk perceptions that influence decisions around beginning or ending a relationship. A plurality of women thought that dating an older man does not bring any adverse consequences, although some thought that older men do not use condoms and may be involved in concurrent partnerships. Many women were less inclined to date same-age or younger men, because they were viewed as being disrespectful and abusive. This study points to the need for more awareness raising about the risks of age-disparate relationships. In addition to these initiatives, there is an urgent need to implement holistic approaches to relationship health, in order to curb intimate partner violence, improve gender equity and make non-age-disparate relationships more attractive.

**Citation:** Beauclair R, Delva W (2013) Is Younger Really Safer? A Qualitative Study of Perceived Risks and Benefits of Age-Disparate Relationships among Women in Cape Town, South Africa. PLoS ONE 8(11): e81748. doi:10.1371/journal.pone.0081748

**Editor:** Paula Braitlein, Indiana University and Moi University, United States of America

**Received:** April 22, 2013; **Accepted:** October 16, 2013; **Published:** November 15, 2013

**Copyright:** © 2013 Beauclair, Delva. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The Flemish Interuniversity Council (VLIR) and the Flemish Scientific Research Fund (FWO) provided financial support for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

\* E-mail: Roxanne.beauclair@gmail.com

### Introduction

Increasingly, epidemiological evidence is showing that age-asymmetric relationships—younger women engaging in sexual relationships with older men—may be at an elevated risk for sexually transmitted infections (STIs), including Human Immunodeficiency Virus (HIV) infection in sub-Saharan Africa [1–6]. The landmark study conducted by Gregson et al., demonstrates that for each year increase in the age difference between young women and their older male partners, the woman's risk of also being HIV positive increases [2]. According to the 2008 South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, the prevalence of infection is significantly higher among young women, with 21.1% of the 20–24 year-old women infected, compared to 5.1% of men infected in the same age category, [7]. The HIV prevalence in men peaks at 25.8% in the 30–34 year-old age group [7].

While the precise mechanism behind this gender discrepancy is incompletely understood, it is thought that a complex interplay of biological, socio-behavioral, and epidemiological factors is responsible for the observed differences in age-stratified HIV prevalence between men and women. Young women may be more vulnerable to HIV infection as a consequence of cervical ectopy (simple columnar epithelium is more susceptible to HIV and other STIs than squamous epithelium), an elevated inflammatory state of the female reproductive tract, and the HIV-susceptibility enhancing effect of injectable hormonal contraceptives [8], which are particularly popular among young women in many countries in sub-Saharan Africa [9]. The age-mixing pattern—how the sexual network connects individuals from different age groups—has also been identified as an important mediator of the spread of STIs [2]. At least two behavioral factors are thought to compound the effects of the biological factors and age-mixing pattern. First, evidence is accumulating to suggest that

condom use is lower in men who engage in sexual relationships with younger women, compared to condom use in other men of the same age [1,10,11]. Second, older partners of young women have been known to sustain multiple concurrent partnerships (MCP) — overlapping sexual partnerships where sexual intercourse with one partner occurs between two acts of intercourse with another partner [2]. Concurrent relationships take place not only with casual partners, but also with main 'long-term' partners. The main partner of men engaging in concurrency is usually older than their casual partners [12], so these men may act as a bridging population, allowing HIV to spread indirectly from older age groups with a higher burden of HIV infections to younger age groups.

Qualitative research has shed light on the factors motivating young women to participate in age-asymmetric relationships in several sub-Saharan African countries. In Leclerc-Madlala's review paper on age-asymmetric relationships, an age-disparate (AD) relationship is defined as one where there is an age difference of five or more years between partners, with the man being older than the woman in the relationship. An inter-generational/cross-generational relationship is a specific type of AD relationship, where the man is ten or more years older than the woman [13]. Often in literature on age mixing, the term 'sugar daddies' is used to refer to men in relationships with younger women, without regard to a specific minimum age difference [14]. It has been suggested that food insecurity and desperation may prompt a young woman to choose an intergenerational relationship [15]. Financial security and material objects that older men are purported to give are some of the more commonly cited incentives for age-asymmetric relationships [13,16–20]. Young women also specify psychosocial benefits such as feelings of love, increased self-esteem and self-confidence [13,17,21], as well as a raised status among peers [18] as reasons for engaging in age-asymmetric relationships. Some young women in Mozambique even enjoyed the freedom and independence that came from these relationships as there is often no expectation of permanency or prolonged affection [16]. Cultural prescriptions and adherence to traditional norms have also been cited as reasons for observing age-asymmetric relationships in sub-Saharan Africa [13].

Despite the mounting evidence about the risks of and motivations for age-asymmetric relationships, it is unclear how women perceive risks of engaging in AD relationships. Establishing how women perceive risks of engaging in AD relationships is necessary in order to inform public health interventions that attempt to limit these relationships and HIV transmission. To this end, we conducted in-depth, open-ended interviews in three different urban, Cape Town communities that are known to have a high prevalence of HIV. The study aimed to explore women's perceptions of risky sexual behaviors and relationship dynamics. Here, we specifically report on their perceived risks of AD and non-age-disparate (non-AD) relationships, as well as women's motivations for engaging in AD relationships. Risk perceptions that influence decisions around starting or breaking up sexual relationships were also explored.

## Materials and Methods

### Ethics Statement

Ethics approval was obtained from Stellenbosch University Health Research Ethics Committee. Written, informed consent was obtained from each participant and community consent was given by community advisory boards and ward councilors in each of the study communities. No compensation was given for any of the interviews.

### Setting

This qualitative study emerged from a previous cross-sectional sexual behavioral survey we conducted from June 2011 to February 2012 in three densely populated urban townships located on the periphery of Cape Town, South Africa. Upon conducting preliminary analysis of the survey data, additional questions began to arise about underlying reasons for risky behaviors including, but not limited to the observed age-mixing patterns. This prompted us to re-engage the three study communities, in March and April 2012, in order to explore relationship dynamics more thoroughly. The details of the antecedent study have been published elsewhere [22].

All three of the communities represent marginalized South African populations where the prevalence of HIV is high. According to the most recently available census data for the City of Cape Town, in all three of the communities more than 40% of the population is unemployed and less than 25% of the population has completed Grade 12 [23]. In two of the three study communities most residents live in informal shack dwellings [23]. In one of the communities participants were predominantly Black, and the other two communities had a combination of Black and Coloured participants. The term 'Coloured' is a South African moniker referring to the racially mixed descendants from Africa, Europe and slave populations of Asia. The three communities were chosen in order to gather a sample of Black and Coloured participants, since people identifying as either of these races often have limited access to financial, health, and educational services, and thus experience the brunt of the HIV epidemic in South Africa.

### Population and Sample

We recruited 23 participants in total, in almost equal numbers from each of the three study communities, using a combination of maximum variation and snowball sampling methodologies. The researcher approached potential female participants, aged 18–65, who were located near the community town hall. Women belonging to a large range of ages were included in the sample in order to gather information about broader norms and beliefs held throughout society about sexual risk behaviors. The purpose of the study was explained to each woman approached, and then, if the woman agreed to participate, a venue and time for the interview was chosen. The number of women refusing to participate was not recorded. At the end of the interview, the women were asked to nominate other women they thought might be interested in participating in the interview. We approached the nominated women using the same method described above. Our sample size was

determined by theoretical saturation. When no new concepts emerged we moved on to the next study community [24].

### Data Collection

The in-depth interviews were conducted by the first author, a female researcher who had postgraduate training in both anthropology and public health. The interviews were held in participants' homes if privacy could be guaranteed. In a few of the cases the participants could not find a time to conduct the interviews at home when family members would not be around, so the researcher secured a safe room in the community hall where the interview was held in complete privacy. A thematic question guide was used to direct the interviews. It was composed of broad themes related to relationship dynamics. The themes were: 'age-disparate relationships', 'starting and breaking up relationships', 'multiple concurrent partners', 'geographical location of partners', 'condom use', 'HIV behavioral interventions', 'alcohol use and sex', and 'partner turnover'. Throughout the interview, participants were allowed to change the direction and scope of the interview in new and meaningful ways. Additionally, the thematic question guide, itself, was not strictly adhered to from one interview to the next. The interviewer made changes in questioning based upon important and unanticipated topics that emerged from previous interviews. Before discussing AD relationships, we asked the participant if she knew what an AD relationship was and if she had heard of the term before. Upon hearing the participant's answer, we provided her with the formal definition and allowed her to respond. The definition we used was: *an age-disparate relationship is one where the male partner is five or more years older than the female partner*. The researcher asked the questions in English and a research assistant was available to translate into isiXhosa or Afrikaans as necessary. Interviews lasted between 30 minutes and 90 minutes. To put the participants at ease, both researchers were female and neither the researcher nor the research assistant were from the study community, and thus, participants could be ensured that personal information would not be shared with the community. Field notes were made after the interview.

### Data Analysis

All interviews were digitally recorded, transcribed verbatim and subsequently translated to English. The data were analyzed using QSR International NVivo 9 software [25]. First, the first author coded all interviews according to the themes in the thematic question guide, developed by the second author. It quickly became apparent that 'age-disparate relationships' and 'starting and breaking up relationships' inspired the richest discussions and from them we gleaned the most novel insights into relationship dynamics. Then a thematic content analysis approach [26] was used to process information within those two themes. The first author read the transcripts line by line and did inductive coding. Themes, such as, 'risks of non-AD relationships' and 'motivations of AD relationships' emerged within these sections of the transcripts. Next, the first author examined text within these themes to see if any new lower-level themes materialized. Meanings and definitions were assigned to these lower-level themes. Both co-authors

discussed the final coding and re-examined transcripts where necessary. The transcripts were not returned to participants for additional feedback primarily due to time constraints and difficulties in re-locating highly mobile participants in informal settlements.

### Results

We interviewed 13 Black and 10 Coloured women, ranging from 20 to 59 years old. Seven of the participants openly discussed being in an AD relationship at the time of the interview. Several of the remaining 16 participants discussed past AD relationships they were involved in as well as ones they knew about at the time of the interview. Coding of these transcripts unearthed several themes related to motivations for participation in AD relationships, as well as perceived risks of AD and non-AD relationships. Finally, in our examination of participants' reported reasons for beginning new relationships and terminating existing ones, we found connections to some of the reported risks of non-AD relationships.

#### Motivations for AD relationships

Overwhelmingly, participants indicated that women were motivated to participate in AD relationships for financial or material benefits. One woman described various basic necessities that were easier to access with the help of an older male partner, "Sometimes let's say you need money for, like, a train ticket, or like, (taxi) fare, or whatever, he can give you money, or if you don't have ... like shoes to wear, like, you know, he could buy you shoes as well" (Aged 40-44, Black). Other women indicated that the need for money, due to impoverishment and desperation, drove women to seek out relationships with older men. One woman stated, "And these young ladies they are also desperate, because the man is going to give the money to the lady, and the lady she's going to pay (them) back" (Aged 25-29, Black). This woman alluded to the fact that some women may be explicitly engaging in transactional sex with older men to try ameliorate their hopeless situation. Another woman offered a similar explanation for her niece's participation in AD relationships, "It comes from poverty, because I was giving (my niece) money, but ... I wasn't giving her enough money ... And sometimes I don't have money because my children are also in private school" (Aged 50-54, Black).

Many participants offered more cynical explanations for why women seek older men. Their articulated views often had an aura of condemnation and the young women were accused of 'using' older men for money:

It's mostly younger women go for married men because sometimes you'll find, like an example, let's say, her neighbor is way older than her, and he's married, and she doesn't have everything like money or anything; life is not that good, so the neighbor, who is married, will offer those things to her. (Aged 25-29, Coloured)

There is some that think the older men, especially the Oupas (grandfathers), they take from them for the money and the house and the car. (Aged 45-49, Coloured)

It's all about money because I don't know of any young woman with an old man who doesn't have money. I don't know such a couple. I don't know. If they have an old man ... Does he have a taxi or a car? It's all about his disability (grant) or his pension. (Aged 45-49, Coloured)

The participants also cited a few psychosocial benefits that result from AD relationships. Specifically, a few women thought that older men could tend to younger women's needs to a greater extent than younger men. They said: "They know how to treat a woman" (Aged 40-44, Black), "An older man will take much better care of a younger woman" (Aged 55-59, Coloured), and "Because older men looks better after you than the younger people" (Aged 25-29, Coloured). The woman from the last quote elaborated by drawing on an example from someone she knew:

Almost like my friend has an older man who could have been her father. Almost like if she didn't have a father figure throughout her childhood life, I assume that is also one of the reasons because maybe she didn't get love and the attention of a father figure; that's why she's looking up to him and he gives her love and attention. (Aged 25-29, Coloured)

Echoing their sentiments, another woman said, "an older guy, he likes to spoil a woman when we ... when you are young, like, he likes to spoil you a lot, and he's got a ... I think ... he's got much more love" (Aged 25-29, Black). Some women believed that an older man's ability to care for a woman and provide love came from maturity that comes with age. As one woman said, "The older guy is more understanding because he's got more life experience" (Aged 40-44, Coloured). Another participant said, "Mostly these younger guys, they want their freedom, they always want to go to the party or whatever, whereas you'd want him to stay at home with you, whereas an older guy, he is more mature and he'll be able to be at home most of the time" (Aged 55-59, Coloured).

Finally, two women expressed the idea that some younger women may choose to engage in AD relationships because of the prestige that is associated with dating older men who have had more opportunities and have acquired more wealth by virtue of their age.

Sometimes the woman will probably be in love with the person, regardless of what the person has. And then with some women, they date the man because he's got his wealth, he's got his money, he's got his car, and he's educated. (Aged 20-24, Black)

Let's say I had an older guy, and you find that this guy ... he's got money, and he's

successful, OK? And most women they do it for those reasons. (Aged 20-24, Black)

The participants were very clear on the motivations young women have for participating in AD relationships, but questions remained about how young women were able to reconcile these motivations with the potential risks that AD relationships may pose for their health.

### Risks involved with AD relationships

A plurality of women, who shared their views about AD relationships, thought that no risks were involved with these relationships. Typical responses included: "From my experience, I don't think there is any risks involved" (Aged 20-24, Black), "I don't see any risk dating someone older" (Aged 20-24, Black), and "No. Like, there is no risk with being with an older man" (Aged 45-49, Coloured). Further probing of these particular women, about how AD relationships might specifically affect the spread of STIs, did not engender any more discussion on the topic. This is, perhaps, not surprising given that almost none of the women had heard of the term 'age-disparate relationships' prior to the interview, nor were they aware of any local health interventions that elucidated risks of AD relationships.

A small minority of women, however, did articulate some concerns with dating older men. They thought that older men refuse to wear condoms. One woman said, "They can't eat a sweet with the wrapper on" (Aged 40-44, Black). Another stated:

(One of my friends) was telling me that 'No, I will not use a condom with my boyfriend, because my boyfriend he doesn't like to use a condom, and then he said that he will never use a condom at his age', whereas he has never used a condom before, you see? ... So when I'm staying and thinking, I think 'Ok, the old guys they say they've never used condoms before, and their grandfathers never used condoms'. So, that means all the older guys they don't use condoms. (Aged 20-24, Black)

A few women explained that the risk to women actually came from the result of non-condom use. "Because sometimes he might not want to use a condom, and ... a mistake could happen and you could, like, fall pregnant" (Aged 40-44, Black). Only two women were specifically concerned that the older men might have more STIs from not wanting to use condoms:

I'll make an example. If I have an older man and I tell him to use a condom and he doesn't want to use a condom, it already tells me something. Maybe this man is not right because he's always supposed to. So ... why don't (he) want to condomize? (He) maybe can have a sick(ness), to make me sick. (Aged 45-49, Coloured)

The younger ones, they've grown up being taught about HIV and STIs, so the older group

they didn't have like a lot of ... they weren't aware of these diseases that we have today, so they are not really ... they're less likely to use condoms. (Aged 20-24, Black)

A small number of women were of the opinion that older men are more inclined to cheat, or in other words, have a concurrent relationship. As one woman stated, "You won't know if that person has other multiple partners ... You'll think that the person really loves you, but not knowing that that person is also busy with other people on the side" (Aged 20-24, Black). Another woman expressed that older men have access to a greater variety of age groups. "That older person have generations of other women" (Aged 25-29, Coloured).

One of the older women in the sample explained that younger women are more naïve and thus not fully aware of all of the other women an older man may be sleeping with and potentially acquiring STIs from:

There are so many things that this man can successfully hide from (her) because of (his) advanced mindset. He can do it ... he can have relationships behind her. This (young girl) is a toddler in front of him. She doesn't know anything. As long as he's giving her love, he can be involved with many other children that are younger than him, and this child can also die of infection. (Aged 50-54, Black)

One woman expressed sentiments about older men being more seasoned when it comes to having sexual relationships, and thus being more likely to have accumulated an STI:

Sometimes dating an older guy, you'll find the guy's more experienced. He's been, like, in this relationship or dating scene for a long time. And then you'll find that he's had many multiple girlfriends, whatever, and then he gets infected by the virus, and then you start dating him, and then you don't use a condom, then he infects you, and then that's how it spreads. (Aged 20-24, Black)

#### Perceived risks of non-AD relationships

Thematic content analysis results suggest that many women might be more concerned about risks associated with non-AD partnerships. Women were primarily concerned that these men might be more abusive and disrespectful to women. As three women remarked: "A guy that is younger, he tends to be more abusive and an older guy will not really lift his hand towards her" (Aged 55-59, Coloured), "Someone younger might be abusive" (Aged 20-24, Black), and "The young ones, they like being abusive and then the older men, he's not like that" (Aged 30-34, Coloured). Another commented, "The younger guys... they're disrespectful, they have no respect and they're very abusive" (Aged 25-29, Black). This woman was particularly concerned about dating same-age men, because she thought they acted as though they had something to prove, "(If) you're dating a younger guy, you will find that the reasons why most of these young guys, like, are abusive is because...that's the

way of claiming respect from you. The older guys they're much nicer, and they don't really, like, tend to be abusive" (Aged 25-29, Black). Another woman deemed same-age men to be more manipulative and emotionally abusive:

No please! I don't like the youngsters. Yhu! ... because they like to play games ... In the meantime they are going to ... remind you that you're an old woman. That is why I don't like that stuff. I know that I'm old, and then now I'm involved with youngsters. And when the time goes on, he's going to remind me every time that "... you know that you're an old woman ... you did date me, and you're too old". (Aged 25-29, Black)

Two women expressed concerns with younger men not providing support in one form or another. One woman described her own personal experience in a non-AD relationship: "I don't trust him because he's younger. I love him, he love me, that was fine ... I'm older than him, I getting sick one of these days. He's not going to look after me; he's going to run, look for a wife his age. Then I will be a loser for the rest of my life" (Aged 45-49, Coloured). Another woman was worried that a non-AD partner would use her. "If you're in a relationship with someone that's closer to your age or younger than you, the guy would want you to buy him a lot of things, but in the relationship where you're with an older guy, it will be more the guy who is able to spoil the woman instead of the woman spoiling the guy" (Aged 40-44, Coloured).

Finally, one woman, who was in an AD relationship with a man nine years older than herself, indicated that men in non-AD relationships might be more likely to engage in multiple concurrent partnerships. She said, "Sometimes when you get a younger guy, they're more into having multiple girlfriends. When you have an older guy, he's more mature, ok?" (Aged 20-24, Black).

#### Influence of risk perceptions on the dissolution of relationships

In the interviews women also discussed typical reasons for deciding to end relationships or seek new ones. The desire to avoid abuse and disrespect in their partnerships was a theme that surfaced in these discussions. When describing why she decided to date her current boyfriend, one woman said, "He is very caring and he would give me money, whereas my child's father didn't give me any money and he would just beat me up and just want to hurt the child" (Aged 55-59, Coloured). Other women articulated similar reasons for deciding to break up relationships based on prior experiences:

I won't normally just end a relationship, but, because my ex-boyfriend was abusive, I just had to end that relationship. (Aged 25-29, Coloured)

If a guy he's cheating on me, and...he's not telling me the truth, and, ja, when the guy is abusing me, I quit. And ... just ... that's all. (Aged 25-29, Black)

Ok the reason why I ended the relationship with the previous guy is because the guy wanted to kill my child. Wanted to stab the child, so I left him and I decided it's just not right and then I met this [other] guy. (Aged 55-59, Coloured)

He was abusive and then if I'd go to work, like lets say I'd go to work the whole week and then Friday come back, I would get paid and he'd take my whole pay. So that's why I ended that other relationship. (Aged 30-34, Coloured)

Maybe the person, like, drinks and then he becomes abusive. And secondly, like, if a person forcefully wants to have sex with you and you don't want to. (Aged 40-44, Black)

Women also considered how respectful their potential partners would be in a relationship. To them, nominal levels of respect were usually required: they just wanted their partners to treat them and their children well. As one woman stated, "He's got to have a very good heart and I've got two kids and he takes care of them and he's very giving towards them" (Aged 30-34, Coloured). Another said, "I'm having a difficult time in terms of the person taking care of me. I'm worried about how he'll take care of me" (Aged 55-59, Coloured). One woman's requirements for initiating a relationship were that her suitor must treat her with courtesy and not be a criminal. "Respect is very important. And then, secondly, the person should have a job, and not go around robbing people" (Aged 25-29, Black). To many these may seem like low standards, but in these communities murders, gender-based violence and thefts are very common [27].

## Discussion

Our qualitative study about Cape Town women's perspectives on relationship dynamics and sexual risk behavior revealed new and useful findings about risk perceptions of AD and non-AD relationships. In accord with other qualitative studies that have looked at motivations for participation in age-asymmetric relationships, our study also found that monetary support and material gain are primary drivers for this behavior [13,16–20]. Moreover, in concordance with recent literature from sub-Saharan Africa that describes how women exercise agency in choice of partner, as well as the duration of the partnership [13,16,18,20], we have also found that some of the participants described 'other women' actively seeking age-asymmetric relationships for psycho-social benefits like, affection, kind-treatment, and prestige. Additionally, some women showed evidence of agency in terminating relationships when they claimed to actively avoid abuse and disrespect. This suggests that many of the women in the study communities are not passive victims entering and staying in relationships out of desperation, but rather they are brokering respect from men whom they deem to be more kind.

Perhaps the most important finding is that a plurality of women did not perceive AD relationships to pose risks of any kind. Even those that perceived some risks did not necessarily

make the connection between those risks and HIV/STI acquisition. The only other qualitative study, to our knowledge, that reported on perceived risks of dating older men, similarly found that both women and men were not concerned with STIs when engaging in these relationships because they judged their partners to be at low risk of having STIs [17].

Our study suggests that women might perceive older men to be a safer choice in comparison to non-AD partners. Women viewed relationships with the latter as particularly risky because they deemed non-AD partners more likely to be abusive. Indeed, women have much to be worried about as 42.3% of men, in a Cape Town-based study, admitted to having perpetrated intimate partner violence (IPV) the last 10 years [28]. Besides the immediate and undesirable threat of violence, qualitative research from South Africa has revealed that women who are victims of IPV believe that they are not in a position to demand condoms from their partners, or ask them to refrain from concurrent relationships [29]. Moreover, IPV has also been linked to other sexual risk indicators, such as meeting a sex partner at a bar/tavern, transactional sex, and recent STI diagnosis [30]. In South Africa, young men under the age of 25 who perpetrate IPV are also more likely to have HIV [31]. Likewise, women who experienced IPV in Rakai, Uganda were at increased risk of HIV infection [32]. It is unclear from our study if the women were aware of these additional risks associated with IPV.

It remains equally uncertain whether or not women's elevated concerns about IPV in relationships with younger men are warranted. In one study that looked at risk factors for committing IPV in South Africa, Gass et al. found that men's age was not a determinant [33]. In studies that looked specifically at age difference as a risk factor for perpetrating or being a victim of IPV, results have been conflicting. Two studies, one conducted in rural South Africa among young women and the other in Denmark among men and women, found an association between having an older partner (>=3 years and >=5 years, respectively) and being a victim of IPV [34,35]. Two studies conducted in the U.S. found no significant association between age difference and female partners being a victim of IPV [36,37]. Finally, a study done in India found that an age difference of 5 years or less between spouses was a risk factor for the female partner being the victim of domestic violence [38]. Irrespective of which age groups truly inflict the most IPV, women's perceptions and expectations are real and their decisions to dissolve relationships and seek new partners are, in part, motivated by their hope to avoid physical and emotional abuse, as we have pointed to in this study. Previous studies have pointed out that many women in sub-Saharan Africa choose to engage in AD relationships for different benefits that often outweigh the perceived risks of engaging in relationships with older men [13,39]. Our study questions whether AD relationships are perceived to be risky at all by a large fraction of women in poor communities with high HIV prevalence around Cape Town. Additionally, it begs the question whether the immediate threat of IPV has a stronger influence on relationship decisions among these women than the inconspicuous and more distant risk of HIV infection and other STIs.

This study conveys the need for a number of key interventions. First, educational and awareness-raising interventions that increase risk perception of AD relationships are currently inadequate and scarce [40,41]. Women need to have all the correct information about the specific types of risk that correspond with particular relationship types if they are to make more informed decisions regarding partner choice. Secondly, women's concern with IPV signals the need for better services that help victims cope with clinical, psychological, social, and legal aspects of IPV. Some of the interventions needed for addressing IPV in these communities are: providing greater care for injuries, mental health, and unwanted pregnancies resulting from rape; delivering more peer support groups for coping with IPV; promoting development of safety plans with referrals to police and health services; and offering legal advice for obtaining protection orders [42]. These services have the potential to increase the social well-being of the affected individuals, as well as decrease anxiety, suicides and alcohol abuse among them [42]. Moreover, services such as these may start to make headway in addressing issues surrounding the substantive use of violence in these communities. These interventions, taken together, may make the less HIV risk-intensive relationships with same-age partners more attractive.

In addition to these interventions, more epidemiological studies need to be conducted in South Africa to explore how risk of IPV and HIV is influenced by partner age differences. Looming questions remain about whether younger or same-age male partners do, in fact, perpetrate more IPV. Additionally, while there is some strong evidence to indicate that increasing age difference puts women at increased risk of being HIV positive, it is still unknown if there is a point at which age-difference stops being a risk factor. The epidemiological studies, to date, have examined this relationship using generalized linear regression models, whereby age difference is treated either as a continuous variable or as categorical variable with arbitrary categories (e.g. >=5 years older), which may mask which ranges of age differences convey the largest risk. Once enough evidence has been gathered, educational programs can attempt to address common misconceptions or truths about whom in their societies tend to be perpetrators of IPV, as well as provide women with information about which 'older' male partners may put them at the most risk of HIV.

As with all qualitative studies, due to our relatively small, non-random sample, we cannot generalize our results to populations outside of urban and disadvantaged Cape Town communities. However, the use of maximum variation sampling procedures means that we were able to capture viewpoints from many types of women in the communities. Furthermore, we determined the sample size through theoretical saturation, which means that no new themes materialized by the time we decided to move on to the next community. By focusing on a limited number of communities only, we were able to gain a better understanding of what motivates women from these communities to choose AD relationships over the alternative, and thus gather context-specific information that will be more useful in developing content for interventions in the study communities. An additional limitation of this study is that some

of the quotes from study participants reference other women's behaviors, which may have limited the accuracy of the experiences they reported. Their opinions could represent hearsay or pluralistic ignorance. Even so, we do not think this was the case most of the time, as nearly one-third of the participants openly admitted to being in an AD relationship at the time of the study and many more admitted to past AD relationships. Moreover, most women did not perceive any grave risks associated with AD relationships or any stigma from participating in them, and therefore our study may have been less affected by social desirability bias. Finally, we are aware that our study is inclusive of age-mature women who may not be at an elevated risk of HIV acquisition through age-disparate relationships. The power dynamics between these women and men are enacted differently than those between younger women and older men. Despite this, we believe the inclusion of older female perspectives is necessary in order to gauge broader norms and beliefs held by society. Indeed, we found that many of the views were cross-generational and cross-racial in our sample.

Our study provides a unique contribution to qualitative literature on the topic of age-asymmetric relationships. To our knowledge, this is one of only two studies [17] to report on the perceived risks of AD and non-AD relationships. Additionally, through our investigation, we were able to generate a new hypothesis: that women's preference for older men may be linked to their avoidance of IPV, thought to be perpetrated more frequently by younger or same-age partners.

## Conclusions

In conclusion, our study of women living in urban Cape Town communities with high HIV prevalence suggests that women may be choosing to engage in AD relationships in part because the alternative—relationships with same-age or younger men—poses the more immediate and severe threat of IPV. In addition to initiatives that raise more awareness about the risks of AD relationships, there is an urgent need to implement holistic approaches to relationship health, in order to curb IPV, improve gender equity and make relationships with peers more attractive.

## Acknowledgements

We wish to thank the women in the three communities for their participation in this study. We are also indebted to our research assistant and translator, Thabisa Maphiri, for countless hours of dedicated fieldwork.

## Author Contributions

Conceived and designed the experiments: RB WD. Performed the experiments: RB. Analyzed the data: RB WD. Contributed reagents/materials/analysis tools: RB WD. Wrote the manuscript: RB. Collected data: RB. Analyzed and interpreted data: RB WD. Edited manuscript: RB WD. Gave final approval for manuscript: RB WD.

## References

- Bankole A, Ahmed FH, Neema S, Ouedraogo C, Konyani S (2007) Knowledge of correct condom use and consistency of use among adolescents in four countries in Sub-Saharan Africa. *Afr J Reprod Health* 11: 197-220.
- Gregson S, Nyamukapa CA, Garnett GP, Mason PR, Zhuwau T et al. (2002) Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *Lancet* 359: 1896-1903.
- Kelly RJ, Gray RH, Sewankambo NK, Serwadda D, Wabwire-Mangen F et al. (2003) Age differences in sexual partners and risk of HIV-1 infection in rural Uganda. *J Acquir Immune Defic Syndr* 32: 446-451.
- Sa Z, Larsen U (2008) Gender inequality increases women's risk of hiv infection in Moshi, Tanzania. *J Biosoc Sci* 40: 505-525.
- Shisana O, Rehle T, Simbayi C, Parker W, Zuma K et al. (2005) South African national HIV prevalence, HIV incidence, behavioural and communication survey, 2005. Cape Town: Human Sciences Research Council.
- Beaulacir R, Kassinjee R, Temmerman M, Welte A, Delva W (2012) Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa. *Eur J Contracept Reprod Health Care* 17: 30-39. doi:10.3109/13625187.2011.648481.
- Shisana O, Rehle T, Simbayi L, Zuma K, Jooste S et al. (2009) South African national HIV prevalence, incidence, behaviour and communication survey 2008: A turning tide among teenagers? Cape Town: Human Sciences Research Council.
- Yi TJ, Shannon B, Prodger J, McKinnon L, Kaul R (2013) Genital immunology and HIV susceptibility in young women. *Am J Reprod Immunol* 69 Suppl 1: 74-79. doi:10.1111/aj.12035.
- Butler AR, Smith JA, Polis CB, Gregson S, Stanton D et al. (2013) Modelling the global competing risks of a potential interaction between injectable hormonal contraception and HIV risk. *AIDS* 27: 105-113. doi:10.1097/QAD.0b000000000000040.
- Luke N (2005) Confronting the 'sugar daddy' stereotype: age and economic asymmetries and risky sexual behavior in urban Kenya. *Int Fam Plan Perspect* 31: 6-14. doi:10.1363/3106050.
- Langeni T (2007) Contextual factors associated with treatment-seeking and high-risk sexual behavior in Botswana among men with symptoms of sexually transmitted infections. *Afr J AIDS Res* 6: 261-269. doi:10.2989/16085900709490422.
- Chopra M, Townsend L, Johnston L, Mathews C, Tomlinson M et al. (2009) Estimating HIV prevalence and risk behaviors among high-risk heterosexual men with multiple sex partners: use of respondent-driven sampling. *J Acquir Immune Defic Syndr* 51: 72-77. doi:10.1097/QAI.0b013e31819907de. PubMed: 19282783.
- Leclerc-Madiala S (2008) Age-disparate and intergenerational sex in southern Africa: the dynamics of hypervulnerability. *AIDS* 22 Suppl 4: S17-S25. doi:10.1097/01.aids.0000343761.77702.04. PubMed: 19033752.
- Kuate-Defo B (2004) Young People's Relationships with Sugar Daddies and Sugar Mummies: What Do We Know and What Do We Need to Know? *Afr J Reprod Health* 8: 13-27. doi:10.2307/3583175. PubMed: 15623116.
- Weiser SD, Leiter K, Bangsberg DR, Butler LM, Percy-de Korte F et al. (2007) Food insufficiency is associated with high-risk sexual behavior among women in Botswana and Swaziland. *PLOS Med* 4: 1589-1597; discussion: 17958460.
- Hawkins K, Price N, Mussa F (2012) Milking the cow: Young women's construction of identity and risk in age-disparate transactional sexual relationships in Maputo, Mozambique. *Global Public Health: An International Journal for Research, Policy and Practice* 4: 169-182.
- Longfield K, Glick A, Waihatika M, Berman J (2004) Relationships between older men and younger women: Implications for STIs/HIV in Kenya. *Stud Fam Plan* 35: 125-34. PubMed: 15260214.
- Luke N (2003) Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa. *Stud Fam Plann* 34: 67-86. doi:10.1111/j.1728-4465.2003.00067.x. PubMed: 12889340.
- Moore AM, Biddlecom AE, Zulu EM (2007) Prevalence and meanings of exchange of money or gifts for sex in unmarried adolescent sexual relationships in sub-Saharan Africa. *Afr J Reprod Health* 11: 44-61. doi:10.2307/25549731. PubMed: 18458736.
- Groes-Green C (2013) "To put men in a bottle": Eroticism, kinship, female power, and transactional sex in Maputo, Mozambique. *Am Ethnologist* 40: 102-117. doi:10.1111/ajem.12008.
- Nkossang J, Rosenthal D (2007) The dynamics of intergenerational sexual relationships: the experience of schoolgirls in Botswana. *Sex Health* 4: 181-187. doi:10.1071/SH08070. PubMed: 17931531.
- Delva W, Beauclair R, Welte A, Vansteelandt S, Hens N et al. (2011) Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol. *BMC Public Health* 11: 616. doi:10.1186/1471-2458-11-616. PubMed: 21810237.
- Town CoC (2001) City of Cape Town-Census 2001. Cape Town: Strategic Development Information.
- Marshall MN (1996) Sampling for qualitative research. *Fam Pract* 13: 522-525. doi:10.1093/fampra/13.6.522. PubMed: 9023528.
- Nvivo qualitative data analysis software: QSR International Pty Ltd.
- Green J, Thorogood N (2004) Qualitative Methods for Health Research. London: SAGE Publications Ltd.
- Mayosi BM, Lawn JE, van Niekerk A, Bradshaw D, Abdoal Karim SS et al. (2012) Health in South Africa: changes and challenges since 2009. *Lancet* 380: 2029-2043. doi:10.1016/S0140-6736(12)61814-5. PubMed: 23201214.
- Abrahams N, Jewkes R, Laubscher R, Hoffman M (2006) Intimate partner violence: prevalence and risk factors for men in Cape Town, South Africa. *Violence Vict* 21: 247-264. doi:10.1891/vi.21.2.247. PubMed: 16642742.
- Fox AM, Jackson SS, Hansen NB, Gasa N, Crewe M et al. (2007) In their own voices: A qualitative study of women's risk for intimate partner violence and HIV in South Africa. *Violence Against Women* 13: 583-602. doi:10.1177/1077801207292909. PubMed: 17515407.
- Pitpitani EV, Kallichman SC, Eaton LA, Cain D, Sikkema KJ et al. (2012) Gender-based violence, alcohol use, and sexual risk among female patrons of drinking venues in Cape Town, South Africa. *J Behav Med* 36: 295-304. PubMed: 22526526.
- Jewkes R, Sikweyiya Y, Morrell R, Dunkle K (2011) The relationship between intimate partner violence, rape and HIV amongst South African men: a cross-sectional study. *PLOS ONE* 6: e24256. doi:10.1371/journal.pone.0024256. PubMed: 21935392.
- Kouyoumdjian FG, Calzavara LM, Bondy SJ, O'Campo P, Serwadda D et al. (2013) Intimate partner violence is associated with incident HIV infection in women in Uganda. *AIDS* 27: 1331-1338. PubMed: 23925380.
- Gass JD, Stein DJ, Williams DR, Seedat S (2011) Gender differences in risk for intimate partner violence among South African adults. *J Interpers Violence* 26: 2764-2789. doi:10.1177/0886260510390960. PubMed: 21156693.
- Jewkes R, Dunkle K, Nduna M, Levin J, Jama N et al. (2006) Factors associated with HIV sero-positivity in young, rural South African men. *Int J Epidemiol* 35: 1455-1460. doi:10.1093/ije/dyl217. PubMed: 17030525.
- Faergemann C, Lauritsen JM, Brink O, Mortensen PB (2010) Do repeat victims of interpersonal violence have different demographic and socioeconomic characters from non-repeat victims of interpersonal violence and the general population? A population-based case-control study. *Scand J Public Health* 38: 524-532. doi:10.1177/1403494810370234. PubMed: 20484309.
- Roberts TA, Auinger P, Klein JD (2006) Predictors of partner abuse in a nationally representative sample of adolescents involved in heterosexual dating relationships. *Violence Vict* 21: 81-89. doi:10.1891/0896-6708.21.1.81. PubMed: 16494134.
- Volpe EM, Hardie TL, Cerulli C, Sommers MS, Morrison-Beedy D (2013) What's age got to do with it? Partner age difference, power, intimate partner violence, and sexual risk in urban adolescents. *J Interpers Violence* 28: 2068-2087. doi:10.1177/0886260512471082. PubMed: 23345572.
- Pandey GK, Dutt D, Banerjee B (2009) Partner and relationship factors in domestic violence: perspectives of women from a slum in Calcutta, India. *J Interpers Violence* 24: 1175-1191. doi:10.1177/0886260508322186. PubMed: 18840848.
- Leclerc-Madiala S (2004) Transactional Modernity, Sex and the Pursuit of Modernity. *Soc Dynam* 2: 1-21.
- Kelly K, Mkhwanazi N, Nkhwashu N, Rapiti R, Mashale R (2012) HIV prevention situation analysis in KwaZulu-Natal, Mpumalanga and Gauteng provinces, South Africa. USAID Sexual HIV Prevention Programme in South Africa (SHPP).
- Hope R (2007) Addressing Cross-Generational Sex: A desk review of research and programs. Population Reference Bureau.
- Joyner K, Mash RJ (2011) The value of intervening for intimate partner violence in South African primary care: project evaluation. *BMJ Open* 1: e000254. doi:10.1136/bmjopen-2011-000254. PubMed: 22146888.

CHAPTER 5: AGE DIFFERENCES BETWEEN SEXUAL PARTNERS, BEHAVIOURAL AND DEMOGRAPHIC CORRELATES, AND HIV INFECTION ON LIKOMA ISLAND, MALAWI

---

---



Photo credit: WorldFish via Visualhunt / CC BY-NC-ND

## 5.1 AIMS OF ANALYSIS

---

The results highlighted in Chapters 3 and 4 generated many hypotheses, however, we decided to go beyond simple visualizations of age-mixing patterns and basic descriptive statistics for our analysis of age-mixing patterns presented in this chapter. The following manuscript was published in *Scientific Reports*. In the manuscript, we aimed to produce a more accurate and nuanced picture of how age-mixing may influence the HIV prevalence in a different population. To that end, we used the LNS data to: 1) Quantify the age-mixing pattern in Likoma Island (**Overall Objective 1**, See Section 1.6); 2) Find associations between relationship characteristics and age differences (**Overall Objectives 2 and 3**); and 3) Explore how three age difference indicators might be related to prevalent HIV infection (**Overall Objective 3**). In this analysis, we employed more advanced statistical techniques to reduce bias introduced by more common modelling strategies.

# SCIENTIFIC REPORTS

OPEN

## Age differences between sexual partners, behavioural and demographic correlates, and HIV infection on Likoma Island, Malawi

Received: 31 May 2016  
Accepted: 11 October 2016  
Published: 02 November 2016

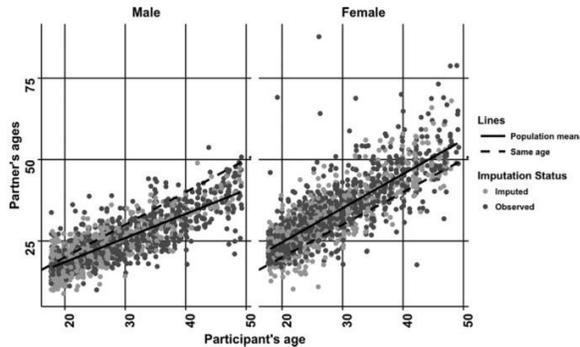
Roxanne Beauclair<sup>1,2</sup>, Stéphane Helleringer<sup>3</sup>, Niel Hens<sup>4,5,6</sup> & Wim Delva<sup>1,2,4,7</sup>

Patterns of age differences between sexual partners – “age-mixing” – may partially explain the magnitude of HIV epidemics in Sub-Saharan Africa. However, evidence of age-disparity as a risk factor for HIV remains mixed. We used data from a socio-centric study of sexual behaviour in Malawi to quantify the age-mixing pattern and to find associations between relationship characteristics and age differences for 1,922 participants. Three age difference measures were explored as predictors of prevalent HIV infection. We found that for each year increase in male participant age, the average age difference with their partners increased by 0.26 years, while among women it remained approximately constant around 5 years. Women in the study had larger within-individual variation in partner ages compared to men. Spousal partnerships and never using a condom during sex were associated with larger age differences in relationships of both men and women. Men who were more than five years younger than their partners had 5.39 times higher odds (95% CI: 0.93–31.24) of being HIV-infected than men 0–4 years older. The relationship between HIV-infection and age-asymmetry may be more complex than previously described. The role that women play in HIV transmission should not be underestimated, particularly in populations with large within-individual variation in partner ages.

Age-mixing results from societal norms for how men and women of a population preferentially choose sexual partners from a certain age. It may impact the spread of HIV<sup>1–3</sup> and partially explain the magnitude of HIV epidemics observed in Sub-Saharan Africa<sup>4</sup>. For instance, a modelling study, based on empirical data from Manicaland, Zimbabwe, implied that a relatively small decrease in the fraction of relationships where the male partner is older than the female partner might reduce the HIV prevalence among young women and their lifetime risk of HIV infection<sup>5</sup>. An ecological study of HIV risk factors in four countries – two with high and two with low HIV prevalence – found that women with older partners were more common in high HIV prevalence settings<sup>6</sup>.

The population-level effect of age-asymmetric relationships may be partially explained by HIV prevalence trends observed in several Sub-Saharan countries. In South Africa and Malawi, for example, HIV prevalence for men tends to peak in older age groups compared to women<sup>7–9</sup>, implying that if women are choosing partners from older age groups where the prevalence is higher than among men in their own age group, they are more likely to choose a partner who has HIV. Additionally, higher sexual risk behaviours have been associated with age-asymmetric relationships. Specifically, men in these relationships use condoms less frequently<sup>10,11</sup>, and have more concurrent partnerships<sup>12–14</sup> than in relationships where men and women are of similar ages. Young women in age-asymmetric relationships also display behaviours and have characteristics which may put them at increased risk for HIV infection: increased reports of STI treatment or symptoms<sup>10,11,15</sup>, more lifetime partners<sup>10</sup>, earlier sexual debut<sup>10</sup>, and higher sex frequency<sup>10</sup>.

<sup>1</sup>International Centre for Reproductive Health, Ghent University, Ghent, Belgium. <sup>2</sup>The South African Department of Science and Technology-National Research Foundation (DST-NRF) Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa. <sup>3</sup>Bloomberg School of Public Health, Johns Hopkins University, Baltimore, USA. <sup>4</sup>Center for Statistics, Hasselt University, Hasselt, Belgium. <sup>5</sup>Centre for Health Economic Research and Modelling Infectious Diseases, Vaccine and Infectious Disease Institute, University of Antwerp, Wilrijk, Belgium. <sup>6</sup>Epidemiology and Social Medicine, University of Antwerp, Wilrijk, Belgium. <sup>7</sup>Rega Institute for Medical Research, KU Leuven, Leuven, Belgium. Correspondence and requests for materials should be addressed to R.B. (email: roxanne.beauclair@gmail.com)



**Figure 1.** Scatter plot of participant ages versus their partners' ages. We used a randomly selected dataset from the 50 datasets we imputed to construct the model. The population mean line resulted from the model predictions.

Observational studies, however, only provide some evidence of the impact of age-asymmetries in relationships on HIV risk. In Zimbabwe, increased age differences between partners raised a participant's odds of being HIV positive<sup>12</sup>. Similarly, in a recent national survey in South Africa, HIV prevalence was higher in those who had age-disparate relationships – defined as a relationship where the partner is five or more years older than the participant – compared to those not in age-disparate relationships<sup>7</sup>. The positive association between age difference and HIV status may, however, only hold for younger women (e.g., less than 25 years old), whereas in older women who have partners older than themselves, HIV or STI risk declines<sup>15,16</sup>. On the other hand, an ecological study in Kenya found no association between the proportion of women who had an age-disparate relationship and HIV prevalence<sup>17</sup>. A study from a large population-based cohort in rural KwaZulu-Natal, South Africa also found no significant association between partner age differences and HIV incidence among women 15–30 years old<sup>18</sup>. Among women in this population who were 30 years or older, risk of HIV acquisition fell as age difference increased<sup>19</sup>.

These conflicting results may be explained by different geographical and cultural contexts. They may also be due to limitations of study designs and measurement of age-mixing. Most studies have focused on the association between age difference during the most recent partnership and HIV risk. Other aspects of age-mixing may also play a role. For example, the maximum age difference that an individual has had with his/her partner(s) may be an important determinant of HIV risk. HIV risk may also be associated with the variation in age differences across the partnerships of an individual<sup>4</sup>. Above and beyond the limitations to measurement and analysis of age-mixing, previous studies have neglected to examine the effects of men choosing older partners on HIV dynamics. In this paper, we used data from the Likoma Network Study in Malawi to test whether these understudied aspects of age-mixing are associated with HIV status.

## Methods

**Data Source.** We used data from the Likoma Network Study (LNS), a sociocentric study of sexual networks and HIV on Likoma Island, Malawi that took place in 2006, 2007/08, and 2013. Data from the LNS were collected in several steps: first, a census of the island's population was collected to obtain a list of potential members of the sexual networks; second, a sexual network survey was conducted with adult members of the population, during which respondents were asked to nominate their recent sexual partners; finally, these nominations were linked to the household census lists, in order to reconstruct maps of the sexual networks that connected inhabitants of the island.

The sexual network survey was administered using Audio Computer-Assisted Self-Interviewing (ACASI), and details of sexual partnerships were recorded for a maximum of five partners in the past three years. Participants were asked to name each of their partners in a recording headset and provide additional identifying information (e.g. location in the village, nicknames, etc.). A partner was defined as someone with whom the participant had had vaginal intercourse within the previous three years. For each partner, questions were asked about the location of their partner, frequency of sexual intercourse, condom use, and additional relationship characteristics.

While the LNS took place in different rounds, the study itself does not follow the strict definition of a cohort study. Some participants that were asked questions in 2006 did not participate in the 2007/2008 version of the study and vice versa. More importantly, some of the sexual behaviour and partner characteristic questions were different from round to round. For example, we do not have partner ages for those participating in 2006. Therefore, we used only data from the 2007/2008 version and treated it as a cross-sectional study. Details of the LNS participant flow have been outlined previously<sup>20</sup>. More details of the LNS study design, sampling, and data collection procedures have been published elsewhere<sup>21</sup>.

Variables	Age difference Unadjusted $\beta$ (95% CI)	
	Men	Women
Relationship is ongoing		
No	Ref	Ref
Yes	0.96 (0.42 – 1.50)	1.28 (0.53 – 2.04)
Partner type		
Spouse	Ref	Ref
Steady partner	-1.56 (-2.32 – -0.80)	-2.19 (-3.16 – -1.22)
Infrequent partner	-0.96 (-1.64 – -0.27)	-1.58 (-2.59 – -0.57)
One-night stand	-1.36 (-2.25 – -0.46)	-1.70 (-3.14 – -0.26)
Last sex with partner		
Within last month	Ref	Ref
Within last year	-0.40 (-1.03 – 0.23)	-0.80 (-1.73 – 0.13)
More than a year ago	-0.81 (-1.44 – -0.19)	-1.67 (-2.52 – -0.82)
Had another partner in relationship		
No	Ref	Ref
Yes	-0.20 (-0.77 – 0.36)	-0.48 (-1.40 – 0.45)
Partner had another partner in relationship		
No	Ref	Ref
Yes	1.02 (0.15 – 1.89)	-0.18 (-1.20 – 0.83)
Yes, suspected	0.41 (-0.25 – 1.06)	0.08 (-0.95 – 1.10)
Do not know	0.46 (-0.32 – 1.24)	0.26 (-0.84 – 1.36)
Sex frequency		
Everyday	Ref	Ref
Several times/week	-0.37 (-1.33 – 0.59)	-0.82 (-2.12 – 0.48)
Once/week	-0.57 (-1.54 – 0.41)	-0.93 (-2.24 – 0.39)
Less than once a week	-0.76 (-1.73 – 0.20)	-0.97 (-2.30 – 0.35)
Ever used a condom in relationship		
No	Ref	Ref
Yes	-0.61 (-1.20 – -0.02)	-1.35 (-2.13 – -0.57)
Residence of partner while in the relationship		
Same village on Likoma	Ref	Ref
Other villages of Likoma	-0.31 (-0.94 – 0.31)	-1.21 (-2.10 – -0.31)
In town on Likoma	-0.07 (-1.06 – 0.91)	-0.36 (-1.78 – 1.05)
Mainland Malawi	-0.52 (-1.42 – 0.39)	-0.57 (-1.73 – 0.59)
Chizumulu	0.21 (-1.33 – 1.76)	-0.08 (-2.11 – 1.96)
Mozambique	0.68 (-1.05 – 2.41)	-0.89 (-3.26 – 1.48)

**Table 1.** Association between age difference and other relationship characteristics. These are pooled estimates from 50 imputed datasets. CI, Confidence Interval.

There were 2009 participants included in this study. However, 87 participants (4.3%) did not report any sexual partners and were thus excluded. Therefore, our secondary analysis includes the 1922 participants from the 2007/2008 round of the LNS that took part in the sexual behaviour survey and reported at least one partner in the past five years. Jointly they reported 3336 relationships, with the average number of partners per participant being 1.74 (Median: 1, IQR: 1–2).

**Measures.** Each study participant self-reported his or her own age. The *age difference* in a given relationship was then defined as the male partner's age minus the female partner's age. Consequently, if the female partner in a relationship was older than her male partner, there would be a negative age difference. For each participant, we calculated the *mean* and *maximum age difference* between an individual and his/her partners. If a participant only reported one relationship then the *mean* and *max age difference* were the same. The *number of partners* variable refers to the number of relationships that the participant reported in the survey. The maximum number of partners a participant could report was five. If the participant had six partners or more, only the most recent five partners would be in the dataset. In our population, only 42 people (2.2%) had five partners in the dataset. In addition, if a respondent had more than five partners and could not report them during the ACASI interview, some of his/her partners may have done so during their own interviews. The relationship may thus have been included in the dataset. Therefore, we believe that the total number of partners was likely not truncated for the vast majority of participants.

We hypothesise that individuals who have relationships with people spanning different birth cohorts may act as 'bridges' for HIV to enter other age groups, and thus we wanted a measure that could capture the range in age

Variables	Crude EBWR (95% CI)	
	Men	Women
Age	1.03 (1.01 – 1.04)	1.02 (0.99 – 1.05)
Highest level of education		
None or Primary	Ref	Ref
Secondary or Tertiary	0.81 (0.60 – 1.09)	0.80 (0.53 – 1.22)
Religion		
Anglican	Ref	Ref
Other	1.13 (0.78 – 1.65)	0.90 (0.52 – 1.55)
Electrified home		
No	Ref	Ref
Yes	0.80 (0.54 – 1.18)	0.92 (0.50 – 1.69)
Owns own home		
No	Ref	Ref
Yes	1.16 (0.77 – 1.74)	1.44 (0.76 – 2.72)
Marital status		
Never Married	Ref	Ref
Divorced, Widowed, Separated	1.00 (0.44 – 2.26)	1.09 (0.60 – 2.01)
Married	1.38 (1.02 – 1.87)	0.89 (0.57 – 1.38)
Condom use		
Never used a condom with partners	Ref	Ref
Used condoms at some point	0.86 (0.59 – 1.25)	1.18 (0.74 – 1.88)
Non-spousal relationships		
Had a non-spousal partner	Ref	Ref
Only had spousal partners	1.05 (0.73 – 1.50)	0.49 (0.20 – 1.20)
Had a once-off relationship		
No	Ref	Ref
Yes	1.14 (0.83 – 1.55)	0.83 (0.50 – 1.37)
Had a concurrent relationship in past 3 years		
No	Ref	Ref
Yes	1.43 (1.06 – 1.92)	1.01 (0.68 – 1.50)
At least 1 partner definitely had another partner		
No	Ref	Ref
Yes	1.28 (0.93 – 1.75)	1.22 (0.82 – 1.81)
Have a partner that currently resides outside of Likoma		
No	Ref	Ref
Yes	0.78 (0.58 – 1.05)	0.83 (0.55 – 1.26)

**Table 2.** Association between bridge width and person-specific characteristics among participants reporting more than one partner in the previous 3 years. These are pooled estimates from 50 imputed datasets. EBWR, Expected Bridge Width Ratio. CI, Confidence Interval.

differences that an individual participant may have. Here we define *bridge width* to be the count of years between the maximum and minimum age difference for an individual. For example, if a woman who was 20 years old at the time of the survey reported two partners: one who was 35 years old and another partner 19 years old at the time of the survey, her *bridge width* would be 16. Finally, the HIV status of participants was determined through rapid HIV testing during LNS data collection<sup>20,21</sup>.

**Statistical Analysis.** Statistical analysis was performed using R version 3.2.2<sup>22</sup>. Our dataset had missing data for many of the key variables used in the analyses (see Supplementary Tables S1–S4). Based on the specific pattern of missingness we observed, we believe that the data were missing at random (MAR). In order to prevent selection bias associated with using complete cases data in an MAR scenario, we used Multivariate Imputation by Chained Equations (MICE), with a Random Forest (RF) algorithm<sup>23</sup> to impute missing values<sup>23–26</sup> (see Supplementary Note S1).

We described the age-mixing pattern by fitting a generalised linear mixed effects model to regress *age of partner* on *age of participant*, with a random intercept for participant, since we have multi-level data. This mixed effects model captures the average partner age, the between-individual variance, and the within-individual variance. The *between-individual variance* is defined as the population-level variance of the participant-specific average age difference between each participant and their partners. The *within-individual variance* refers to how much the age differences between a participant and his or her partners vary around that participant's average age difference with his or her partners. The model allowed for heteroskedastic residual variance using a power

Variables	Men aOR (95% CI)			Women aOR (95% CI)		
	Mean AD Model	Maximum AD Model	BW Model	Mean AD Model	Maximum AD Model	BW Model
Number of partners	1.28 (0.98 – 1.67)	1.29 (0.98 – 1.69)	1.17 (0.84 – 1.63)	1.51 (1.22 – 1.87)	1.47 (1.19 – 1.82)	1.48 (1.16 – 1.90)
Age of participant	1.01 (0.70 – 1.45)	1.01 (0.70 – 1.45)	1.00 (0.70 – 1.44)	1.04 (0.84 – 1.29)	1.04 (0.83 – 1.29)	1.04 (0.84 – 1.29)
Age of participant squared	1.01 (0.99 – 1.04)	1.02 (0.99 – 1.04)	1.01 (0.99 – 1.04)	1.01 (0.99 – 1.03)	1.01 (0.99 – 1.03)	1.01 (0.99 – 1.03)
Age of participant cubed	1.00 (1.00 – 1.00)	1.00 (1.00 – 1.00)	1.00 (1.00 – 1.00)	1.00 (1.00 – 1.00)	1.00 (1.00 – 1.00)	1.00 (1.00 – 1.00)
Mean age difference						
0–4 years older	Ref	–	–	Ref	–	–
6 or more years younger	5.39 (0.93 – 31.24)	–	–	2.31 (0.60 – 8.98)	–	–
1–5 years younger	0.89 (0.13 – 6.00)	–	–	1.20 (0.55 – 2.64)	–	–
5–9 years older	1.43 (0.68 – 2.98)	–	–	1.04 (0.62 – 1.76)	–	–
10 or more years older	0.80 (0.30 – 2.10)	–	–	1.40 (0.78 – 2.51)	–	–
Maximum age difference						
0–4 years older	–	Ref	–	–	Ref	–
6 or more years younger	–	4.08 (0.40 – 41.16)	–	–	2.51 (0.63 – 9.99)	–
1–5 years younger	–	1.79 (0.36 – 8.83)	–	–	1.53 (0.68 – 3.42)	–
5–9 years older	–	1.58 (0.71 – 3.51)	–	–	1.18 (0.69 – 2.02)	–
10 or more years older	–	1.16 (0.47 – 2.88)	–	–	1.51 (0.86 – 2.65)	–
Bridge width	–	–	1.04 (0.95 – 1.13)	–	–	1.00 (0.94 – 1.06)

**Table 3.** Association between age difference measures and HIV status. These are pooled estimates from 50 imputed datasets. aOR, adjusted Odds Ratio. CI, Confidence Interval. AD, Age Difference. BW, Bridge Width.

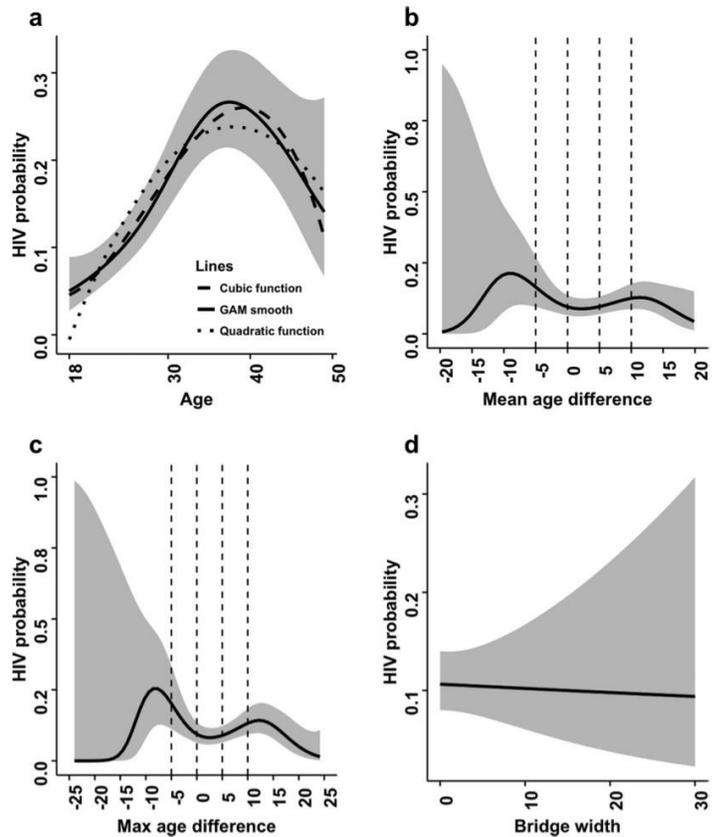
variance function structure, since preliminary explorations of the data suggested that the variance of the residuals increases with increasing age of the participant.

Next we used generalized linear mixed effects models with a random intercept on participants to determine if *age difference* is associated with relationship characteristics in men and women. We also calculated the median *bridge widths* for men and women among participants who reported more than one partner. Then we used this sub-population of participants and examined the relationship between person-specific characteristics and *bridge width* – our outcome variable. To do this we used negative binomial regression, with the imputed datasets, stratified by gender. Negative binomial regression is useful when the outcome of interest represents overdispersed count data that is bounded by zero. We exponentiated the coefficients produced by the negative binomial model in order to obtain an Expected Bridge Width Ratio (EBWR). In both the *bridge width* and *age difference* models we were only concerned with finding predictors of these age difference constructs using marginal associations, rather than building causal models that adjust for potential confounders.

Finally, we examined the association between our three age difference measures and prevalent HIV using logistic regression to calculate adjusted Odds Ratios (aORs) and 95% confidence intervals (95% CIs). We categorized *mean age difference* and *max age difference* (6 or more years younger/1–5 years younger/0–4 years older/5–9 years older/10 or more years older) while keeping *bridge width* as a linear term in the models. All models were adjusted for *number of partners*, as well as squared and cubic polynomial terms for *age of participant*. The analyses were stratified by gender.

Choosing cut-off points for the categories in our analytical models has a three-fold purpose: 1. It is convenient and makes it easy to compare results across studies and settings (some prominent examples include<sup>6,13,14</sup>), 2. It also allows us to directly compare the men and women in our study population, and 3. It permits us to pool estimates from models repeated on all 50 imputed datasets. However, we also wanted to qualitatively describe the relationship between our age difference measures and the probability of being HIV infected, without forcing a linear or parametric relationship. We suspected the relationship between age difference and HIV was more complex than our parametric analytical models allowed us to observe. Therefore, we randomly selected one of the imputed datasets and constructed Generalised Additive Models (GAMs)<sup>27</sup> to use as the basis for plots that show the graphical relationship between the age difference measures and the probability of having HIV. These semiparametric models used P-spline smoothers for the age difference measure, and adjust for the same terms as our logistic regression models described above. Unfortunately, the estimates from GAMs cannot be pooled as required for multiply imputed data.

**Ethical Approval.** Institutional Review Boards at the University of Malawi College of Medicine and at the University of Pennsylvania approved the LNS. Institutional Review Board approval to conduct this secondary analysis was obtained from the Stellenbosch University Health Research Ethics Committee (IRB0005239). The methods were carried out in accordance with the approved guidelines. Informed consent was obtained from all participants in the LNS. For this secondary analysis no participants were contacted and no new data were collected for this analysis.

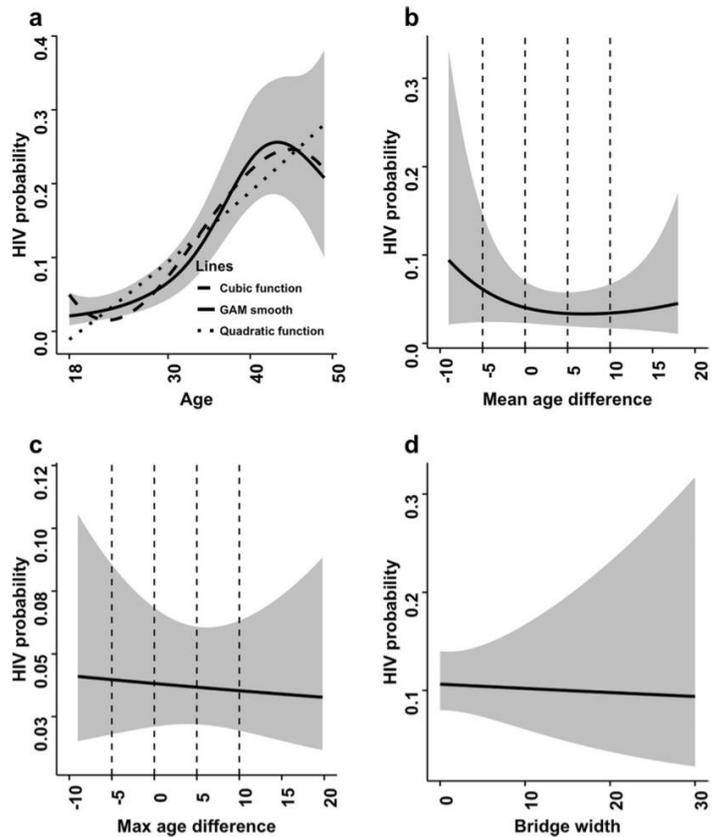


**Figure 2.** Smoothed plots showing the probability of being HIV-positive among women. The model in panel (a) is univariate and used to demonstrate the functional form that age should take in the regression models. The confidence bands presented result from the GAM. Models in panels (b) to (d) adjust for age, age squared, age cubed, and total number of partners in the past 5 years. For comparison purposes, the chosen cut-points to categorize the age difference measure from the parametric analysis are indicated by vertical dashed lines.

## Results

**Age-mixing pattern.** Our survey sample was comprised of 1068 women and 854 men reporting on 1648 and 1688 relationships, respectively. Figure 1 illustrates the age-mixing pattern, with the predictions from the model superimposed. In both genders there was a positive linear relationship between age of participant and partners' ages. For each year increase in male participant age, the average age difference with his partners increased by 0.26 years. The average age difference between female participants and their partners remained approximately constant around 5 years. In 18 year olds, the between-individual variance in partner ages was slightly higher for women compared to men (1.66 [95% CI: 0.33–8.33] vs. 1.21 [95% CI: 0.63–2.33], respectively). In the same aged participants the within-individual variance was 9.16 (95% CI: 7.69–10.92) for men and 17.05 (95% CI: 13.62–21.35) for women.

**Association between relationship and participant characteristics with age difference.** Table 1 shows that there were smaller age differences in relationships where the last sexual episode took place more than a year ago compared to within the last month. This was true for men ( $\beta = -0.81$ , 95% CI:  $-1.44$  to  $-0.19$ ) and



**Figure 3.** Smoothed plots showing the probability of being HIV-positive among men. The model in panel (a) is univariate and used to demonstrate the functional form that age should take in the regression models. The confidence bands presented result from the GAM. Models in panels (b) to (d) adjust for age, age squared, age cubed, and total number of partners in the past 5 years. For comparison purposes, the chosen cut-points to categorize the age difference measure from the parametric analysis are indicated by vertical dashed lines.

women ( $\beta = -1.67$ , 95% CI:  $-2.52$  to  $-0.82$ ). In relationships of male participants, the average age difference was 1.56 years smaller with 'steady partners' compared to spouses, while in women the age difference was 2.19 years smaller. The other non-spousal partnerships were also associated with smaller age differences. Those that had ever used a condom with their partner also had reduced age differences: almost 1 year less for men and 1.35 years for women. Women in relationships with men living in other villages on Likoma at the time of the relationship had, on average, a 1.21 years (95% CI:  $-2.1$  to  $-0.31$ ) smaller age difference than those with partners in the same village. Among men who believed that their partner had another partner outside of the relationship, there were, on average, larger age differences ( $\beta = 1.02$ , 95% CI:  $0.15$ – $1.89$ ) compared to men who thought their partners never had any other partners while with them. Finally, in both men and women, relationships that were ongoing at the time of the survey had larger age differences than those that had been terminated.

In our sample, 875 participants reported more than one partner in the past three years and had a median bridge width of 4 years (IQR: 1.5–6.5). Specifically for men the median was 3.5 years (IQR: 1.1–5.9) and for women it was 5 years (IQR: 2.5–7.5). Out of all the participants who had more than one partner, only 12 (1.4%) had a bridge

width equal to zero, i.e., all their partners had the same age difference. Table 2 illustrates that among men reporting more than one partner, age, marital status, and concurrency were all predictive of larger bridge widths. Older men were more likely to have larger bridge widths. The expected bridge width reported by men increased by 3% for each year increase in age (EBRW: 1.03, 95% CI: 1.01–1.04). Larger bridge widths were also expected among married men (EBWR: 1.38, 95% CI: 1.02–1.87) and those with a concurrent relationship in the past three years (EBWR: 1.43, 95% CI: 1.06–1.92). Among women, the only predictor of larger bridge widths was age (EBWR: 1.02, 95% CI: 0.99–1.05).

**Relationship between age difference measures and HIV status.** In the parametric analysis of the association between our age difference measures and HIV status (Table 3), the number of partners a participant had was associated with prevalent HIV in all female models. However, it was only marginally associated with increased likelihood of HIV in male mean and max age difference models. Men who were, on average, six or more years younger than their partners, had five times higher odds of being HIV positive (95% CI: 0.93–31.24) than men who were 0–4 years older than their partners. Regarding women, none of the models demonstrated a relationship between our categorized age difference constructs and HIV status.

Figures 2 and 3 depict the results of our gender-stratified semiparametric analysis. The relationship between participant age and the probability of being HIV positive appears to most closely resemble a cubic function (Figs 2a and 3a), and is thus the reason why we used that functional form in our regression models. HIV probability increases slowly until about 28 years in women and 33 years in men, and then increases more rapidly until 38 years and 43 years, respectively. Thereafter, the probability starts to decline. Of particular note is the bimodal relationship between mean and max age difference and HIV probability observed in women (Fig. 2b,c). Women who are older than their partners tend to have an increased probability of being HIV positive, but the risk decreases the closer the partner is to her own age. Then the risk for women starts to gradually increase again as their partners become older. When the male partner is 12 or more years older, the risk declines again. A more pronounced, but qualitatively similar pattern is observed for max age difference. When examining the mean age difference plot for men (Fig. 3b), a markedly different relationship is observed. Male participants who were younger than their partners had an increased probability of being HIV positive, and that risk declined as their ages became more similar. Figures 2 and 3 should be interpreted with caution, particularly at extreme values for the age difference measures since the confidence intervals are large due to the sparse number of observations.

## Discussion

Our study quantified the age-mixing pattern among sexually active adults on Likoma Island in Malawi, and explored several ways of analysing the relationship between age difference and prevalent HIV. We considered several age difference measures, as well as parametric and semiparametric regression techniques and discovered that the relationship between HIV and age difference may be more complex than previously described.

We found an age-mixing pattern of increasing average age differences as men get older. For women there is a constant average age difference of approximately five years. This pattern is similar to that observed in a study of adolescents and young adults in Cape Town, South Africa<sup>11</sup>. Additionally, a study that took place in rural KwaZulu-Natal, South Africa also found that age differences in current partners of women who were aged 18–30 years old were approximately four years older. From ages 30–80, the age gaps slowly decreased to almost zero<sup>28</sup>. In our study, we only examined participants 18–49 years old, so we do not know if we would witness a similar decrease if older women had been included in the study. It is possible that in the South African context of extremely high HIV prevalence – where men get tested and treated less often than women – as women become widows in their middle-aged years they begin to take new partners that are closer to their own age.

The findings from our parametric regression analysis that examined the associations between three age difference measures and HIV status are in apparent agreement with recent studies conducted in South Africa. Harling *et al.* did not find evidence that among women 15–30 years old, age differences were associated with increased hazard of HIV acquisition<sup>18</sup>. Likewise, Street *et al.* also could not reject the null hypothesis that women who were in age-disparate relationships would have the same HIV incidence as those in similar-age relationships<sup>14</sup>. However, our results from the semiparametric analysis indicate that this may not be the full story. In particular, women may have some increased risk of being HIV positive if their male partners are older than them by approximately 2–12 years, but then may be slightly protected if their partners are more than 12 years older. The analysis conducted by Street *et al.* produced similar, albeit non-significant, trends: the HIV incidence among women in steady sexual partnerships was 5.78 per 100 person-years with partners who were 0–4 years older, 7.50 with partners 5–9 years older, and 3.67 with partners 10 or more years older<sup>14</sup>. While in our study the associations between mean/max age differences and HIV status in women were not statistically significant, we do believe there is in fact a bimodal association that is meaningful and relevant for HIV transmission dynamics.

Our age-mixing study goes beyond describing mean age differences, and breaks down the variation of age differences into a between-individual and within-individual component. The relatively large within-individual variation in partner ages for young women means that there are opportunities to acquire HIV from men in one age group and then transmit to men in another age group. The potential for transmission between age groups is particularly high because in both men and women in our study, individuals who had larger age differences were more likely to be in spousal relationships, never use a condom during sex, and have had sex in the month prior to the survey. Moreover, men who had a partner whom they thought had a simultaneous relationship, also had larger age differences.

Largely absent from the discourse about how age differences influence HIV epidemiology is how men are affected by age-asymmetric relationships. In our study population, it was not uncommon for men to have had some relationships where the female partner was older than them. Importantly, our study suggests that these men may be at increased risk of having HIV. While not specifically examining age asymmetry and HIV status, a recent

study of youth in Eastern Cape, South Africa has also demonstrated that older partner age was associated with curable STIs among boys<sup>39</sup>. Additionally, Gregson *et al.* found that men who chose partners of similar age were more likely to be HIV infected<sup>12</sup>.

One crucial limitation of our study is that the extreme values for the age difference measures in our semiparametric analysis had very large confidence intervals, and therefore, these areas should not be over-interpreted. Secondly, the cross-sectional nature of our study means that we could not determine causality between age difference measures and HIV status. Using complex, semiparametric models requires large sample sizes and so the link between age difference measures and incident HIV could be explored through simulations in a data-driven model world when conducting a large prospective study is not feasible.

We likely also had too few observations in our analysis of the association between bridge widths and HIV status, since calculating a bridge width requires reporting at least two relationships. Less than half of the participants reported more than one. However, we believe that bridge width has the potential to be an important individual-level indicator of variability in partner ages, and should be considered in future studies of age-asymmetry and HIV transmission dynamics, particularly in contexts where individuals are inclined to report more than one partner. In a previous modelling study, it was found that small changes to the variance in the distribution of age differences is enough to increase the basic reproductive number for HIV transmission<sup>4</sup>. The same study also concluded that if the variance in the distribution was large, the mean age difference for the population did not matter much for transmission. We found that men who were married and those who reported a concurrent partner in the past three years were inclined to have larger bridge widths. Moreover, there was also a relationship between older ages and increasing bridge widths. Taking these associations together with the observation that the highest prevalence of HIV among men occurs between the ages of 35 and 45 years implies that there is great HIV transmission potential between different age groups.

Despite these limitations, we believe that our study boasts several strengths. First, we employed two different methods for examining associations between age difference measures and prevalent HIV. The use of GAMs allowed us to build a flexible model that did not impose a specific parametric response function<sup>27</sup>. We also constructed parametric models, which produced no evidence for an association between age difference and HIV status. The benefit of constructing these two types of models, fitting the same data, allowed us to investigate the implication of applying traditional, parametric models to data that require more flexible, nuanced models. Traditionally in analyses of age difference and HIV, if age difference was not treated as a linear term in models, then it was categorized. These arbitrary categories may not adequately capture the true underlying risk patterns<sup>30</sup>. An additional advantage of smooth plots is the ability to visually compare risk among any two or more values of the age difference measures, instead of just comparing to one reference group. For example, we were able to see that women who had a maximum age difference of eight years with their partners had considerably higher risk than those who were at most three years younger than their partners.

Another strength of our study was that the survey data for this secondary analysis was collected using ACASI. ACASI is believed to help elicit more reporting of sensitive behaviours in African contexts<sup>31,32</sup> and to reduce social desirability bias<sup>33</sup>. Thus we believe the reports of relationship characteristics and sexual behaviours in our study were likely more accurate compared to those from surveys where face-to-face interviewing was conducted.

The study of age-asymmetry and HIV transmission dynamics is over two decades old, but the degree to which we understand how choice in partner age influences HIV is still in its infancy. In particular, the role women may play in HIV transmission pathways that connect men from different age groups is still largely unexplored. Our study provides evidence suggesting that the way in which we analyse age-asymmetry may be overly simplistic and require more flexible models to determine the real and complex relationship that age difference, age, and number of partners have to HIV. We believe a crucial step in this line of research is to employ similar analytical techniques to evaluate how different age difference measures – particularly bridge widths – relate to HIV incidence in either large cohorts from different settings or simulation studies modelled after real-world contexts. Though we found that women have large within-individual variation in age difference, we still lack the empirical evidence base to confirm or reject the notion that variation in partner age is a critical driver of individual-level risk of prevalent and incident HIV infection.

## References

- Anderson, R. M., May, R. M., Ng, T. W. & Rowley, J. T. Age-Dependent Choice of Sexual Partners and the Transmission Dynamics of HIV in Sub-Saharan Africa. *Philosophical Transactions: Biological Sciences* 336, 135–155 (1992).
- Bershteyn, A., Klein, D. J. & Eckhoff, P. A. Age-dependent partnering and the HIV transmission chain: a microsimulation analysis. *Journal of the Royal Society, Interface/the Royal Society* 10, 20130613, doi:10.1098/rsif.2013.0613 (2013).
- Garnett, G. P. & Anderson, R. M. Sexually transmitted diseases and sexual behavior: insights from mathematical models. *The Journal of infectious diseases*. 174 Suppl 2, S150–S161 (1996).
- d'Albis, H., Augeraud-Veron, E., Djemai, E. & Ducrot, A. The dispersion of age differences between partners and the asymptotic dynamics of the HIV epidemic. *Journal of biological dynamics*, 6, 695–717, doi: 10.1080/17513758.2012.688146 (2012).
- Hallett, T. B., Gregson, S., Lewis, J. J., Lopman, B. A. & Garnett, G. P. Behaviour change in generalised HIV epidemics: impact of reducing cross-generational sex and delaying age at sexual debut. *Sexually transmitted infections*. 83 Suppl 1, i50–i54, doi: sti.2006.02360610.1136/sti.2006.023606 (2007).
- Chapman, R. *et al.* Do behavioural differences help to explain variations in HIV prevalence in adolescents in sub-Saharan Africa? *Tropical Medicine and International Health*. 15, 554–566 (2010).
- Shisana, O. *et al.* South African National HIV Prevalence, Incidence and Behaviour Survey, 2012 (HSRC Press, Cape Town, 2014).
- Pettifor, A. E. *et al.* Young people's sexual health in South Africa: HIV prevalence and sexual behaviours from a nationally representative household survey. *Aids*. 19, 1525–1534, doi: 00002030-200509230-00012 (2005).
- Office, N. S. & Macro, I. Malawi Demographic and Health Survey 2010 (2011).
- Morrison-Beedy, D., Xia, Y. & Passmore, D. Sexual risk factors for partner age discordance in adolescent girls and their male partners. *Journal of clinical nursing*. 22, 3289–3299, doi: 10.1111/jocn.12408 (2013).
- Beauchair, R., Kassinjee, R., Temmerman, M., Welte, A. & Delva, W. Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa. *The European journal of contraception & reproductive health care: the official journal of the European Society of Contraception*. 17, 30–39, doi: 10.3109/13625187.2011.644841 (2012).

12. Gregson, S. *et al.* Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *Lancet* **359**, 1896–1903, doi: 10.1016/S0140-6736(02)08780-9 (2002).
13. Maughan-Brown, B., Kenyon, C. & Lurie, M. N. Partner Age Differences and Concurrency in South Africa: Implications for HIV-Infection Risk Among Young Women. *AIDS and behavior*, **18**, 2469–2476, doi: 10.1007/s10461-014-0828-6 (2014).
14. Street, R. A., Reddy, T. & Ramjee, G. The generational effect on age disparate partnerships and the risk for human immunodeficiency virus and sexually transmitted infections acquisition. *International journal of STD & AIDS*, **27**, 746–752, doi: 10.1177/0956462415592325 (2016).
15. Kraut-Becher, J. R. & Aral, S. o. Patterns of age mixing and sexually transmitted infections. *International journal of STD & AIDS*, **17**, 378–383 (2006).
16. Kelly, R. J. *et al.* Age differences in sexual partners and risk of HIV-1 infection in rural Uganda. *Journal of acquired immune deficiency syndromes*, **32**, 446–451 (2003).
17. Kenyon, C. R., Vu, L. & Menten, J. & Maughan-Brown, B. Male Circumcision and Sexual Risk Behaviors May Contribute to Considerable Ethnic Disparities in HIV Prevalence in Kenya: An Ecological Analysis. *PloS one*, **9**, e106230, doi: 10.1371/journal.pone.0106230 (2014).
18. Harling, G. *et al.* Do Age-Disparate Relationships Drive HIV Incidence in Young Women? Evidence from a Population Cohort in Rural KwaZulu-Natal, South Africa. *Journal of acquired immune deficiency syndromes*, **66**, 443–451, doi: 10.1097/QAI.0000000000000198 (2014).
19. Harling, G., Newell, M. L., Tanser, F. & Barnighausen, T. Partner Age-Disparity and HIV Incidence Risk for Older Women in Rural South Africa. *AIDS and behavior*, **19**, 1317–1326, doi: 10.1007/s10461-014-0952-3 (2015).
20. HELLERINGER, S., MKANDAWIRE, J., KALLANI-PHIRI, L. & KOHLER, H. P. Cohort Profile: The Likoma Network Study (LNS). *International journal of epidemiology*, **43**, 545–557, doi: 10.1093/ije/dyt001 (2014).
21. HELLERINGER, S., KOHLER, H. P., CHIMBIRI, A., CHATONDA, P. & MKANDAWIRE, J. The Likoma Network Study: Context, data collection, and initial results. *Demographic research*, **21**, 427–468, doi: 10.4054/DemRes.2009.21.15 (2009).
22. R. A language and environment for statistical computing. (R Foundation for Statistical Computing, Vienna, Austria, 2014).
23. Shah, A. D., Bartlett, J. D., Carpenter, J., Nicholas, O. & Hemingway, H. Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study. *American journal of epidemiology*, **179**, 764–774 (2014).
24. Penone, C. *et al.* Imputation of missing data in life-history trait datasets: which approach performs best? *Methods in Ecology and Evolution*, **5**, 961–970 (2014).
25. Stekhoven, D. J. & Bühlmann, P. MissForest - nonparametric missing value imputation for mixed-type data. *Bioinformatics*, **28**, 112–118 (2012).
26. van Buuren, S. & Groothuis-Oudshoorn, K. MICE: Multivariate Imputation by Chained Equations in R. *Journal of statistical software*, **45**, 1–67 (2011).
27. Shiboski, S. C. Generalized Additive Models for Current Status Data. *Lifetime data analysis*, **4**, 29–50 (1998).
28. Ott, M. Q., Barnighausen, T., Tanser, F., Lurie, M. & Newell, M. L. Age-gaps in sexual partnerships: seeing beyond 'sugar daddies'. *Aids*, **25**, 861–869 (2011).
29. O'Leary, A. *et al.* Associations between Psychosocial Factors and Incidence of Sexually Transmitted Disease Among South African Adolescents. *Sexually transmitted diseases*, **42**, 135–139 (2015).
30. Bennette, C. & Vickers, A. Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents. *BMC medical research methodology*, **12**, 21, doi: 10.1186/1471-2288-12-21 (2012).
31. Dolezal, C. *et al.* A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Archives of sexual behavior*, **41**, 401–410, doi: 10.1007/s10508-011-9769-6 (2012).
32. Hewett, P., Mensch, B. & Erulkar, A. Consistency in the reporting of sexual behavior by adolescent girls in Kenya: a comparison of interviewing methods. *Sexually Transmitted Infections*, **80**, i43–i48 (2004).
33. Beauclair, R. *et al.* Evaluating Audio Computer Assisted Self-Interviews in Urban South African Communities: Evidence for Good Suitability and Reduced Social Desirability Bias of a Cross-sectional Survey on Sexual Behaviour. *BMC medical research methodology*, **13**, doi: 10.1186/1471-2288-13-11 (2013).

### Acknowledgements

We would like to thank Jonathan Dushoff at McMaster University and Guy Harling at Harvard University for their ongoing conversations about the theoretical ways in which age-mixing may impact the spread of HIV. Their advice and feedback helped us to formulate our secondary analysis study design. RB and WD received funding from the Research Foundation - Flanders (FWO). NH acknowledges support from the University of Antwerp scientific chair in Evidence-Based Vaccinology, financed in 2009–2016 by an unrestricted donation from Pfizer.

### Author Contributions

The concept of this study was conceived by R.B., S.H., N.H. and W.D. S.H. was the principal investigator of the LNS. R.B. conducted the statistical analysis and took a lead in writing the manuscript. R.B., S.H., N.H. and W.D. contributed to discussion of the results, and editing of the final manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Beauclair, R. *et al.* Age differences between sexual partners, behavioural and demographic correlates, and HIV infection on Likoma Island, Malawi. *Sci. Rep.* **6**, 36121; doi: 10.1038/srep36121 (2016).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016

**Supplementary Table S1. Comparison of participant characteristics by HIV missing status.** All continuous values are reported with median and inter-quartile range, Med (IQR), while categories are reported in percentages, n (%).

Variables	HIV status missing	
	No n = 1,474	Yes n = 448
Age	27.0 (22.0 - 34.0)	29.5 (23.0 - 36.0)
Gender		
Male	612 (41.5%)	242 (54.0%)
Female	862 (58.5%)	206 (46.0%)
Highest level of education		
None or Primary	863 (58.5%)	238 (53.1%)
Secondary or Tertiary	561 (38.1%)	195 (43.5%)
Missing	50 (3.4%)	15 (3.3%)
Religion		
Anglican	1,182 (80.2%)	330 (73.7%)
Other	263 (17.8%)	105 (23.4%)
Missing	29 (2.0%)	13 (2.9%)
Electrified home		
No	1,204 (81.7%)	348 (77.7%)
Yes	265 (18.0%)	97 (21.7%)
Missing	5 (0.3%)	3 (0.7%)
Owns own home		
No	201 (13.6%)	91 (20.3%)
Yes	1,268 (86.0%)	354 (79.0%)
Missing	5 (0.3%)	3 (0.7%)
Marital status		
Never Married	490 (33.2%)	132 (29.5%)
Divorced, Widowed, Separated	145 (9.8%)	38 (8.5%)
Married	803 (54.5%)	264 (58.9%)
Missing	36 (2.4%)	14 (3.1%)
Number of partners	1.0 (1.0 - 2.0)	1.0 (1.0 - 2.0)

**Supplementary Table S2. Comparison of participant characteristics for those with any missing partner age versus those with all known partner ages.** All continuous values are reported with median and inter-quartile range, Med (IQR), while categories are reported in percentages, n (%).

Variables	Partner age missing	
	A partner age is missing n = 896	No ages missing n = 1,026
Age	25.0 (21.0 – 32.0)	29.0 (23.0 – 36.0)
Gender		
Male	445 (49.7%)	409 (39.9%)
Female	451 (50.3%)	617 (60.1%)
Highest level of education		
None or Primary	472 (52.7%)	629 (61.3%)
Secondary or Tertiary	394 (44.0%)	362 (35.3%)
Missing	30 (3.3%)	35 (3.4%)
Religion		
Anglican	716 (79.9%)	796 (77.6%)
Other	167 (18.6%)	201 (19.6%)
Missing	13 (1.5%)	29 (2.8%)
Electrified home		
No	735 (82.0%)	817 (79.6%)
Yes	158 (17.6%)	204 (19.9%)
Missing	3 (0.3%)	5 (0.5%)
Owens own home		
No	132 (14.7%)	160 (15.6%)
Yes	761 (84.9%)	861 (83.9%)
Missing	3 (0.3%)	5 (0.5%)
Marital status		
Never Married	414 (46.2%)	208 (20.3%)
Divorced, Widowed, Separated	127 (14.2%)	56 (5.5%)
Married	334 (37.3%)	733 (71.4%)
Missing	21 (2.3%)	29 (2.8%)
Number of partners	2.0 (2.0 – 3.0)	1.0 (1.0 – 1.0)

**Supplementary Table S3. Comparison of relationship characteristics by participants' HIV missing status.** All continuous values are reported with median and inter-quartile range, Med (IQR), while categories are reported in percentages, n (%).

Variables	HIV status missing	
	No n = 2,601	Yes n = 735
Partner is alive		
Alive	1,456 (56.0%)	395 (53.7%)
Dead	145 (5.6%)	44 (6.0%)
Do not know	8 (0.3%)	0 (0.0%)
Missing	992 (38.1%)	296 (40.3%)
Relationship is ongoing		
No	1,608 (61.8%)	436 (59.3%)
Yes	981 (37.7%)	292 (39.7%)
Missing	12 (0.5%)	7 (1.0%)
Partner type		
Spouse	1,025 (39.4%)	311 (42.3%)
Steady partner	583 (22.4%)	152 (20.7%)
Infrequent partner	722 (27.8%)	175 (23.8%)
One-night stand	260 (10.0%)	93 (12.7%)
Missing	11 (0.4%)	4 (0.5%)
Last sex with partner		
Within last month	951 (36.6%)	323 (43.9%)
Within last year	608 (23.4%)	174 (23.7%)
More than a year ago	1,005 (38.6%)	223 (30.3%)
Missing	37 (1.4%)	15 (2.0%)
Had another partner in relationship		
No	2,004 (77.0%)	557 (75.8%)
Yes	584 (22.5%)	173 (23.5%)
Missing	13 (0.5%)	5 (0.7%)
Partner had another partner in relationship		
No	713 (27.4%)	210 (28.6%)
Yes	569 (21.9%)	152 (20.7%)
Yes, suspected	852 (32.8%)	244 (33.2%)
Do not know	425 (16.3%)	116 (15.8%)
Missing	42 (1.6%)	13 (1.8%)
Sex frequency		
Everyday	227 (8.7%)	76 (10.3%)
Several times/week	763 (29.3%)	224 (30.5%)
Once/week	752 (28.9%)	179 (24.4%)
Less than once a week	815 (31.3%)	245 (33.3%)
Missing	44 (1.7%)	11 (1.5%)
Ever used a condom in relationship		
No	956 (36.8%)	252 (34.3%)
Yes	1,500 (57.7%)	373 (50.7%)
Missing	145 (5.6%)	110 (15.0%)
Residence of partner while in the relationship		
Same village on Likoma	1,166 (44.8%)	298 (40.5%)
Other villages of Likoma	614 (23.6%)	155 (21.1%)
In town on Likoma	262 (10.1%)	67 (9.1%)
Mainland Malawi	377 (14.5%)	157 (21.4%)
Chizumulu	91 (3.5%)	24 (3.3%)
Mozambique	73 (2.8%)	24 (3.3%)
Missing	18 (0.7%)	10 (1.4%)

**Supplementary Table S4. Comparison of variables for those with missing partner age versus observed partner age.** All continuous values are reported with median and inter-quartile range, Med (IQR), while categories are reported in percentages, n (%).

Variables	Partner age missing	
	No n = 1,997	Yes n = 1,339
Partner is alive		
Alive	883 (44.2%)	968 (72.3%)
Dead	51 (2.6%)	138 (10.3%)
Do not know	2 (0.1%)	6 (0.4%)
Missing	1,061 (53.1%)	227 (17.0%)
Relationship is ongoing		
No	929 (46.5%)	1,115 (83.3%)
Yes	1,055 (52.8%)	218 (16.3%)
Missing	13 (0.7%)	6 (0.4%)
Partner type		
Spouse	1,081 (54.1%)	255 (19.0%)
Steady partner	384 (19.2%)	351 (26.2%)
Infrequent partner	392 (19.6%)	505 (37.7%)
One-night stand	131 (6.6%)	222 (16.6%)
Missing	9 (0.5%)	6 (0.4%)
Last sex with partner		
Within last month	969 (48.5%)	305 (22.8%)
Within last year	423 (21.2%)	359 (26.8%)
More than a year ago	573 (28.7%)	655 (48.9%)
Missing	32 (1.6%)	20 (1.5%)
Had another partner in relationship		
No	1,534 (76.8%)	1,027 (76.7%)
Yes	455 (22.8%)	302 (22.6%)
Missing	8 (0.4%)	10 (0.7%)
Partner had another partner in relationship		
No	538 (26.9%)	385 (28.8%)
Yes	422 (21.1%)	299 (22.3%)
Yes, suspected	630 (31.5%)	466 (34.8%)
Do not know	385 (19.3%)	156 (11.7%)
Missing	22 (1.1%)	33 (2.5%)
Sex frequency		
Everyday	212 (10.6%)	91 (6.8%)
Several times/week	647 (32.4%)	340 (25.4%)
Once/week	559 (28.0%)	372 (27.8%)
Less than once a week	544 (27.2%)	516 (38.5%)
Missing	35 (1.8%)	20 (1.5%)
Ever used a condom in relationship		
No	748 (37.5%)	460 (34.4%)
Yes	1,085 (54.3%)	788 (58.8%)
Missing	164 (8.2%)	91 (6.8%)
Residence of partner while in the relationship		
Same village on Likoma	1,140 (57.1%)	324 (24.2%)
Other villages of Likoma	495 (24.8%)	274 (20.5%)
In town on Likoma	153 (7.7%)	176 (13.1%)
Mainland Malawi	128 (6.4%)	406 (30.3%)
Chizumulu	43 (2.2%)	72 (5.4%)
Mozambique	23 (1.2%)	74 (5.5%)
Missing	15 (0.8%)	13 (1.0%)

## Supplementary Note S1. Rationale for choice in imputation methods

Multivariate Imputation by Chained Equation (MICE) methods create multiple datasets using other variables in the original dataset to predict missing values. Analysts can then use the same regression models on each dataset and produce pooled estimates, which take into account uncertainty in the imputations [1] and lead to larger confidence intervals. In order to reduce simulation error, that may result from imputing too few datasets we imputed 50 datasets [1]. The main advantage of using the Random Forest (RF) algorithm with MICE imputation is that the algorithm makes few assumptions about the data and therefore, it is ideal for imputing datasets for which there are nonlinearities and interactions between key variables [2-4].

### References

1. van Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate Imputation by Chained Equations in R. *Journal of statistical software*. 2011;45(3):1-67.
2. Penone C, Davidson AD, Shoemaker KT, DiMarco M, Rondinini C, Brooks TM, et al. Imputation of missing data in life-history trait datasets: which approach performs best? *Methods in Ecology and Evolution*. 2014;5(9):961-70.
3. Shah AD, Bartlett JD, Carpenter J, Nicholas O, Hemingway H. Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study. *American journal of epidemiology*. 2014;179(6):764-74.
4. Stekhoven DJ, Bühlmann P. MissForest - nonparametric missing value imputation for mixed-type data. *Bioinformatics*. 2012;28(1):112-8.

CHAPTER 6: PARTNER AGE DIFFERENCES AND  
ASSOCIATED SEXUAL RISK BEHAVIOURS AMONG  
ADOLESCENT GIRLS AND YOUNG WOMEN IN A CASH  
TRANSFER PROGRAMME FOR SCHOOLING IN  
MALAWI

---

---



Photo credit: IMs Bildarkiv / Photo Archive via Visualhunt.com / CC BY-NC-ND

## 6.1 AIMS OF ANALYSIS

---

In Chapter 3, we presented the finding that young women in the CAPS population, who had attended more classes were less commonly involved in ADR. Other researchers have since then published findings from various structural interventions aimed at keeping young girls in school in order to improve their long-term economic prospects and also reduce the risk of HIV acquisition. The SIHR study (See Section 2.2.3) was one such study, which found that girls who received the intervention had a lower prevalence of HIV, as well as, a lower prevalence of relationships with men 25 years or older. This finding was intriguing and it prompted us to investigate whether the choice of younger partners by girls in the intervention groups of this study may have partially influenced the observed HIV prevalences. We hypothesized that girls receiving income via cash transfers, would have smaller age differences, because they would not have strong financial motivations for pursuing those age-asymmetric relationships. We also wanted to determine what association, if any, existed between age differences and relationship-level characteristics that affect HIV transmission. These two aims fall within the scope of my **overall objective 3** (See Section 1.6). The following manuscript has been submitted to *BMC Public Health* and is currently under review.

## **Partner age differences and associated sexual risk behaviours among adolescent girls and young women in a cash transfer programme for schooling in Malawi**

Roxanne Beauclair <sup>\*, 1, 2</sup>, Jonathan Dushoff <sup>3</sup>, Wim Delva <sup>1, 2, 4, 5</sup>

<sup>1</sup> International Centre for Reproductive Health, Ghent University, De Pintelaan 185, 9000 Gent, Belgium

<sup>2</sup> The South African Department of Science and Technology-National Research Foundation (DST-NRF) Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, 19 Jonkershoek Road, Stellenbosch 7600, South Africa

<sup>3</sup> Department of Biology, McMaster University, 1280 Main St. West, Hamilton, Ontario, L8S 4K1, Canada

<sup>4</sup> Centre for Statistics, Hasselt University, Martelarenlaan 42, BE3500 Hasselt, Belgium

<sup>5</sup> Rega Institute for Medical Research, KU Leuven, Herestraat 49, 3000 Leuven, Belgium

\* Corresponding author

Email addresses:

RB: roxanne.beauclair@gmail.com

JD: dushoff@mcmaster.ca

WD: wim.delva@ugent.be

## **Abstract:**

**Background:** Age disparities in sexual relationships have been proposed as a key risk factor for HIV transmission in Sub-Saharan Africa, but evidence remains inconclusive. The SIHR study, a cluster randomised trial of a cash transfer programme in Malawi, found that young women in the intervention groups were less likely to have had a sexual partner aged 25 or older, and less likely to test positive for HIV and HSV-2 at follow-up compared to control groups. We examined the hypotheses that girls in the intervention groups had smaller age differences than control groups and that large age differences were associated with relationship-level HIV transmission risk factors: inconsistent condom use, sex frequency, and relationship duration.

**Methods:** We conducted an analysis of schoolgirls in the SIHR study aged 13-22 at baseline (n = 2907). We investigated the effects of study arm, trial stage and participant age on age differences in sexual relationships using linear mixed-effects models. Cumulative-link mixed-effects models were used to estimate the effect of relationship age difference on condom use and sex frequency, and a Cox proportional hazard model was used to estimate the effect of relationship age difference on relationship duration. We controlled for the girl's age, number of partners, study group and study round.

**Results:** Girls receiving cash transfers, on average, had smaller age differences in relationships compared to controls, though the estimated difference was not statistically significant (-0.43 years; 95% CI: -1.03, 0.17). The older the participant was, the smaller her age differences (-0.67 per 4-year increase in age; 95% CI: -0.99, -0.35). Among controls, after the cash transfers had ended the average age difference was 0.82 years larger than during the intervention (95% CI: 0.43, 1.21), suggesting a possible indirect effect of the study on behaviour in the community as a whole. Across treatment groups, larger age differences in relationships were associated with lower levels of condom use, more frequent sex, and longer relationship durations.

**Conclusions:** Cash-transfer programmes may prevent HIV transmission in part by encouraging young women to form age-similar relationships, which are characterised by increased condom use and reduced sex frequency. The benefits of these programmes may extend to those who are not directly receiving the cash.

**Key words:** Malawi, Sexual Risk Behaviour, Age-disparate Relationships, Age-Mixing, Southern Africa

## Background

“Age-mixing” and its effect on HIV transmission dynamics is an important research area in the quest for an AIDS-free generation (1,2). Age-mixing patterns – how individuals in a population choose partners with regards to age – influence HIV transmission in modelling studies (3–5) and may partially explain variation in epidemic size in sub-Saharan Africa (6). Evidence of increased individual-level risk is mixed, with studies showing a positive association of relationship age differences with HIV prevalence (7–10), though not with HIV incidence (11–13).

There are at least a few causal pathways by which younger partners may confer lower risk for HIV infection. First, age-similar young male partners are more likely to have been sexually active for a shorter period of time, and thus be less likely to be HIV positive (8,14–16). Second, older male partners are more likely to be in a concurrent relationship (7,13,17), which increases the likelihood young women will be sexually active with a man when he experiences acute HIV infection. Finally, age-asymmetric relationships have been reported to result in inconsistent condom use (18–22) and higher frequencies of sex (20,21). This could be due to the gender-power imbalances that may be more likely in age-disparate relationships, and which result in less ability of younger women to negotiate when and how sex occurs (21,23).

Given these causal mechanisms, large relationship age differences are likely to increase women’s risk of HIV infection. We believe the failure to detect a consistent relationship between age differences and HIV infection by previous studies may be partially related to how data has been analysed. Until recently, empirical age-mixing studies have analysed observational data in two primary ways. The first way involves categorizing the most recent, or primary relationship into 2 or 3 categories: typically, the relationship is “non-age-disparate”, “age-disparate” (where the male partner is 5 or more years older than the female partner), and sometimes “inter-generational” (the male partner is 10 or more years older) (e.g. 12,13). In the second type of analysis linear correlations between relationship age differences (as a continuous variable), and HIV infection risk or prevalence are analysed (e.g. 11). In both types of analyses relationships in which the women are older than their partners are typically discarded. However, recent studies suggest that the effect of relationship age difference on prevalent HIV infection is nonlinear, as well as age-specific (24–26).

An example demonstrating this complexity was recently published from a study on Likoma Island in Malawi (24). They found that women who were approximately 2–12 years younger than their partners had higher probabilities of being HIV positive than women in more age-similar partnerships, but those who were more than 12 years younger than their partners had lower probabilities of being HIV positive (24). Additionally, a phylogenetic study of individuals from a high-HIV-prevalence population in South Africa found that girls and young women aged 15-25 years old were most likely infected by men who were in their 30s, while those same men were mostly infected by women aged 25-40 years (25). These studies indicate a need to move beyond the current paradigm and utilize analytical techniques that allow for flexible, nonlinear relationships between variables. Arbitrary categorizations and forced linear relationships may mask the true underlying risk patterns (27–30).

Recently, several structural and behavioural interventions have aimed to increase school attendance in young African girls, improve long-term economic outcomes, and thus indirectly reduce HIV transmission (9,31–33). A cash transfer programme in Zomba, Malawi – the Schooling, Income, and Health Risk (SIHR) study – found that the

prevalence of HIV and HSV-2 was lower among 13-22 year-old girls in the intervention arm, compared to the control group, at their 18-month follow-up (34). The intervention arm also had a lower prevalence of relationships with men 25 years or older. Some studies have suggested that young women are often motivated to engage in relationships with older men because the men may provide them with pocket money or gifts (35–38), money for school fees (39), and food (40). The findings from the SIHR study do not necessarily mean that the lower partner ages in the intervention groups directly caused a reduction in HIV transmission, but they do suggest the hypothesis that young women receiving cash transfers may have had less motivation to choose older partners who would put them at increased risk of HIV infection, through the mechanisms identified above.

In the Malawian context, where HIV peaks in men at older ages than in women (14,16), interventions that inhibit relationships with older men may be key to reducing incidence among adolescent girls and younger women. Using flexible nonlinear models, and publicly available data from the SIHR study, we assessed evidence for the hypothesis that smaller age differences in the intervention groups was a driver of their lower observed STI prevalences. To do this, we investigated the effect of the intervention on age differences, as well as, examined the association between age differences and relationship-level characteristics that affect HIV transmission risk: condom use, sex frequency, and relationship duration.

## Methods

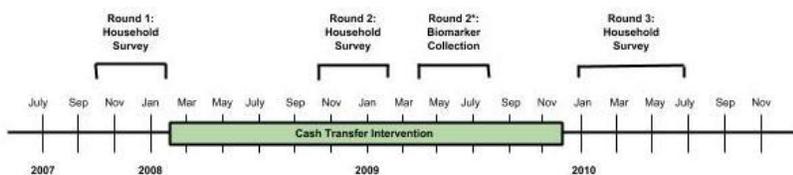
### *Study design and data source*

We conducted a secondary analysis of the SIHR study data that comes from Zomba, Malawi. Zomba district includes one of the four largest cities in Malawi (Zomba City), in addition to a large rural population (41). At the time of the SIHR study, it was primarily agricultural, with most people participating in subsistence farming (42), and only 6% of the adult population having a formal income as of 2008 (41). HIV/AIDS was, and still is, one of the largest public health problems facing the district (43). The SIHR study was a cluster randomised trial of a cash transfer programme administered to girls 13-22 years old (34). Only aspects relevant to the present analysis will be discussed here.

Enumeration Areas (EAs) in the Zomba district were randomly sampled ( $n = 176$ ) from urban, rural, and near-urban areas. Households with at least one never-married girl aged 13-22 years old were included in the study. Eighty-eight of the EAs were randomised to the intervention group, and the other 88 to the control group. Girls in both groups were classified as schoolgirls if they were enrolled in school at baseline, and otherwise as dropouts. Schoolgirls within intervention EAs were randomly assigned to receive: unconditional cash transfers (UCTs); conditional cash transfers (CCTs), paid only if they attended school at least 80% of the days school was in session; or no cash transfers (Spillovers). The study design included Spillovers in order to see if there would be an indirect effect of the trial on girls in intervention areas who did not receive cash. One hypothesis is that the spillover girls may also change their behaviour, because they see other girls (i.e. the ones receiving cash transfers) changing their behaviour (44). The main difference between Controls and Spillovers, was that no one in the control EAs received cash. Please see Baird et al. for the trial profile from the original study (34). The principle investigators of the SIHR study also did a comparison of baseline characteristics for the different study groups to ensure that randomization was effective

(34, 42). Cash was administered monthly to both households and girls in the intervention groups.

Data were publicly available for three rounds of data collection: Round 1 (baseline) took place before allocation of intervention assignment; Round 2 (R2) was during the intervention; and Round 3 (R3) was after the end of the intervention (Figure 1; (45)). The survey was administered in two parts. The first part was administered to the heads of households and obtained information on household characteristics. The second part was administered to the participating girls and solicited responses about their health, dating patterns, and social networks.



**Figure 1. Overview of SIHR study intervention and data collection rounds.** \*HIV data collected during Round 2 biomarker collection were not made publicly available, and therefore, not used in our analysis.

### *Participants*

At baseline 2907 schoolgirls were sampled and interviewed. Our focal outcome was the age difference in a relationship, which was only recorded during R2 and R3. In each of those rounds girls could report up to 3 partners that they had in the past 12 months. Since our focus was age differences, we studied only schoolgirls who reported sexual relationships in the past 12 months at either R2 or R3. It is a shortcoming of our study that we do not have partner ages in R1, and therefore cannot control for random differences in place before the study started, nor can we detect changes in behaviour between R1 and R2. Despite this, we believe the changes that occur between R2 (during the intervention) and R3 (after the intervention) are still informative and provide insight into what policy makers and researchers might expect at the conclusion of a similar intervention. The majority of girls did not report a relationship in either round, and thus, we were left with 1108 schoolgirls who reported on a total of 1491 relationships in R2 and R3 combined. Relationships were the unit of observation in our analyses.

### *Statistical analysis*

Our analyses focus on four outcome variables potentially affecting HIV transmission risk: 1. *age difference* (continuous, one-year increments); 2. *condom use* (ordinal, Never/Inconsistent/Every time); 3. *sex frequency* (ordinal, Once or twice only/Less than twice per month/A couple times per month/1-3 times per week/4 or more times per week); and 4. *relationship duration* (continuous, one-week increments). We chose these variables because the probability that a girl will become infected with HIV by a sex partner is largely determined by the frequency of unprotected sex within that relationship. Inconsistent condom use, high numbers of sex acts in a given period of time, and long relationship durations, are relationship characteristics that may increase the opportunities for HIV transmission, and they may act as causal mediators in the relationship between age differences and HIV risk.

Age difference was calculated by subtracting the girl's age from the age of her partner. This variable was used as the outcome in our first model, and then as an explanatory variable in subsequent models. Relationship duration was determined by asking girls how long ago the relationship began and how long ago the relationship ended. Relationships that were ongoing at the time of the survey were treated as having right-censored durations.

To investigate the effects of study arm, girl age, and survey round (i.e. during versus after the intervention) on relationship age difference, we fitted a linear mixed-effects regression model. Prior to constructing our model, we chose three planned contrasts for the study arm comparisons: pooled UCTs/CCTs versus controls, UCTs versus CCTs, and spillovers versus controls. We defined the pooled UCTs/CCTs to be the average of the two effects. Our primary interest was the contrast between the combined intervention groups and controls, because we wanted to determine if the intervention would have an effect on choosing partner ages, regardless of conditionality. We hypothesized that intervention group girls would have smaller age differences with their partners compared to the control group during R2 when monthly cash transfers were taking place. We believed they would be less motivated to choose an older partner who might provide them with pocket money, because they would have their own source of income. Additionally, we decided to compare spillovers versus controls to see if there was an indirect effect (cash was not received by spillovers) of being in an intervention EA. We also included a study arm-by-survey round interaction term in our model.

In two additional models, we treated condom use and sex frequency as response variables and used cumulative-link mixed effects models (CLMMs) to assess the relationship between each of these outcomes and age difference. CLMMs are commonly used when the outcome of interest is ordinal. The underlying assumption is that the effect of the covariates on the log odds of moving to a higher response level is the same at each response level. To assess the effect of age difference on relationship duration, while accounting for censorship, we used a Cox proportional hazards model.

Since sexual behaviours of young women are expected to change through time as they get older, and this effect could be conflated with effects of intervention timing, we adjusted for the participant's age in our models. We also controlled for the number of partners in the previous 12 months, study group and study round, in the models where condom use, sex frequency, and relationship duration were outcomes, because each of these variables could have potentially confounded the relationship between age differences and the sexual behaviours. In all of the mixed-effects models we had random intercepts for the EA and girl. In the Cox model, we used robust sandwich variance estimators to take into account clustering at the level of the EA. Survey sample weights

were not used in these analyses since we were not concerned with estimating population-level descriptive statistics (46).

As discussed above, flexible nonlinear models allow us to investigate a variety of risk patterns when age differences and age are continuous covariates in models. To determine the functional form of these variables in the models for condom use, sex frequency and relationship duration, we first fitted generalized additive models (GAMs) with penalized regression smoothers for the continuous variables, while adjusting for model covariates. GAMs are semiparametric models that allow continuous variables to ‘speak for themselves’ without imposing a specific form (47). The estimated Effective Degrees of Freedom (EDF) for the continuous term was then rounded to the nearest integer and used to determine how many degrees of freedom (DF) to use for our spline terms in the final models. In the CLMM for condom use we implemented a spline for age difference with DF equal to 3 and kept age as a linear term (equivalent to a spline with 1 DF). In the CLMM for sex frequency, the spline terms for age difference and age of girl had DFs equal to 4 and 2, respectively. In the Cox model age difference had 2 DFs and age was kept as a linear term. Splines are difficult to interpret using model coefficients; we therefore used effects plots to visualize how the outcomes of interest varied as a function of the predictors. All statistical analyses were performed with R version 3.3.1 (48).

## Results

### *Description of key variables*

Overall, there were 1491 relationships reported: 541 in R2 (36.3%) and 950 in R3 (63.7%). Controls reported 783 relationships (52.5%), spillovers 315 relationships (21.1%), CCTs 275 relationships (18.4%), and UCTs 118 relationships (7.9%).

Figure 2a shows that in all study groups and rounds there are larger proportions of the girls using condoms “Never” or “Inconsistently” than using “Every time”. The fraction of girls who reported using condoms “Every time” decreased from R2 to R3 among all study groups; this was most pronounced among the controls. In all study groups the fraction of girls who reported sex 1-3 times per week increased from R2 to R3, with the greatest jump among the UCTs: from 18.0% to 42.3% (Figure 2b). Sex four or more times per week was relatively uncommon in all study groups. Figure 2c shows that average age differences increased from R2 to R3 in all groups. The average relationship duration increased from R2 to R3 for all study groups (Figure 2d). It is probable that part of this increase across all study groups may be related to the fact that some relationships are with the same partner in R2 and R3, and thus, the relationships would have been ongoing for a longer period of time by R3. There is an especially large increase in mean duration for spillovers and controls. We hypothesize that these groups either stopped partaking in short-term relationships, or refrained from reporting them if they happened to have many of them.

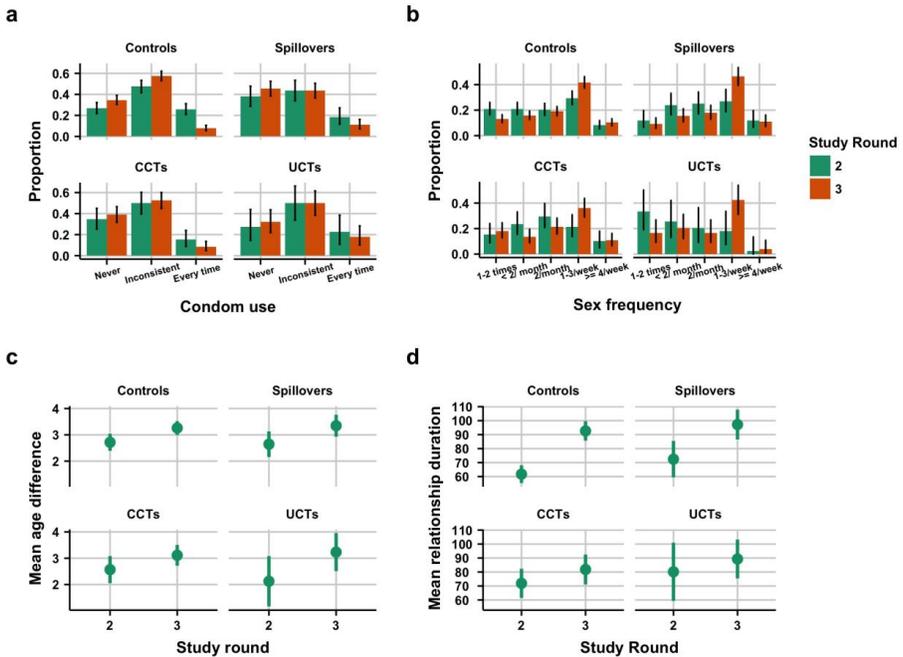
### *Effects of SIHR on relationship age differences*

Results of our linear mixed-effects model (Figure 3) show that for approximately every four years (2 SD = 3.96 years) increase in age of the girl, the average age difference

decreased by 0.67 years (95% CI: -0.99, -0.35). The average age difference was 0.82 years (95% CI: 0.43, 1.21) larger among control girls reporting in the post-intervention period (R3) compared to during the time of the intervention (R2). The observed effect of study group on average age difference was in the hypothesized direction with CCTs and UCTs having a smaller difference than control girls (-0.43 years; 95% CI: -1.03, 0.17), though the effect was not statistically significant. The overall pattern of effects can also be seen in the Supplementary Figure 1, Additional File 1.

*Condom use model*

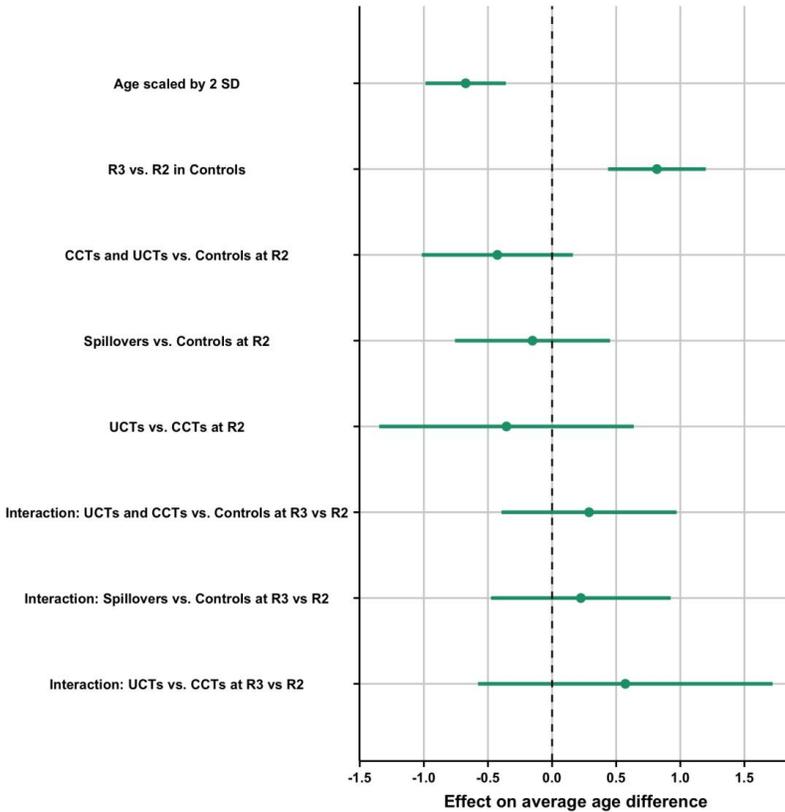
Figure 4a shows that the probability of using a condom “Every time” decreased as the absolute value of age difference between the girl and her partner increased. Figure 4b shows overall condom use decreased from equal-age partnerships through to an age difference of about 7 years. Patterns outside this range cannot be inferred clearly, because the CIs are large. Figure 4c shows that the odds of higher versus lower categories of condom use decreased by 18% (POR: 0.82; 95% CI: 0.64, 1.06) for approximately every four years increase in girls’ age. In R3 compared to R2, girls had 35% (POR: 0.65; 95% CI: 0.51, 0.83) lower odds of reporting more condom use. In spillovers versus controls, girls had lower odds (POR: 0.68; 95% CI: 0.48, 0.97) of higher levels of condom use. A similar, but non-significant, effect was observed in CCTs versus controls.



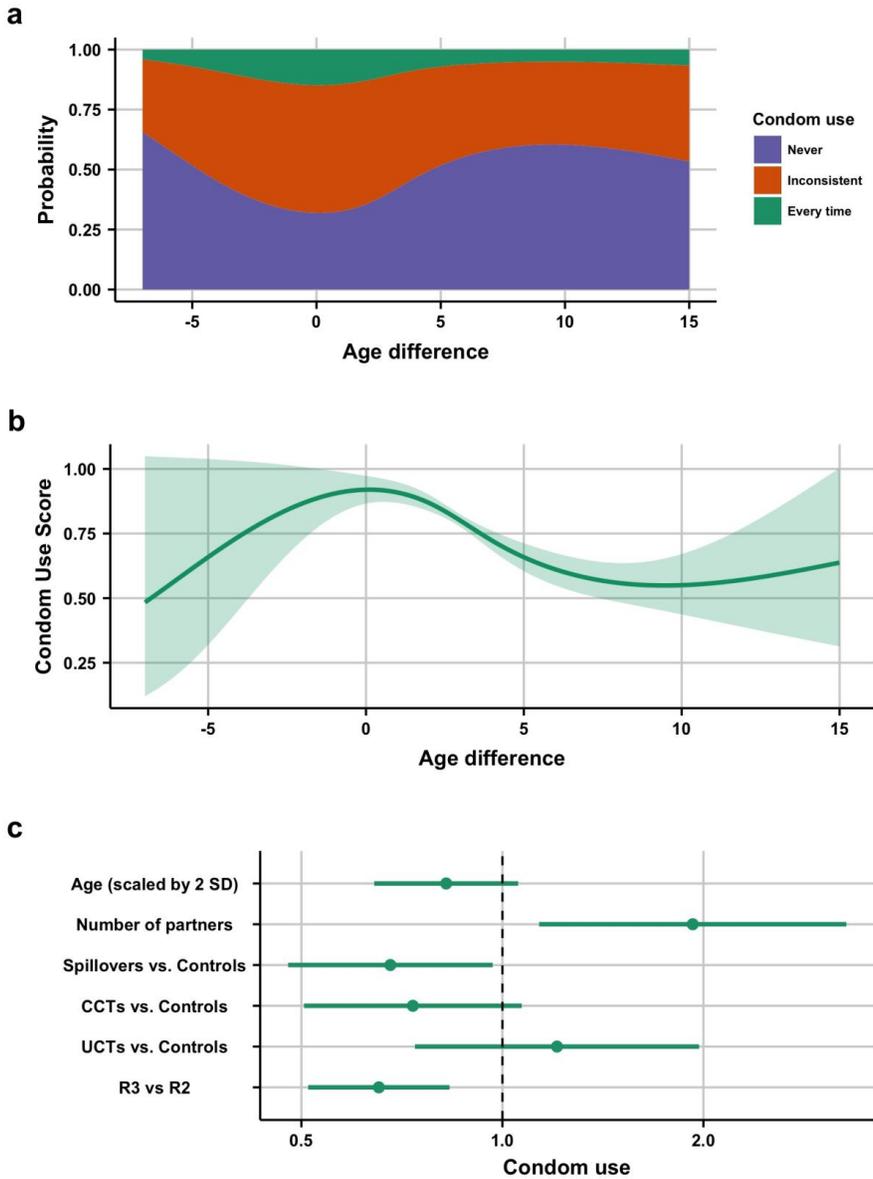
**Figure 2. Summary statistics for relationship characteristics, by study group and round.** The panels contain summaries for: a. condom use (n = 1491); b. sex frequency (n = 1490); c. age difference (n = 1364); and d. relationship duration (n = 1256).

*Sex frequency model*

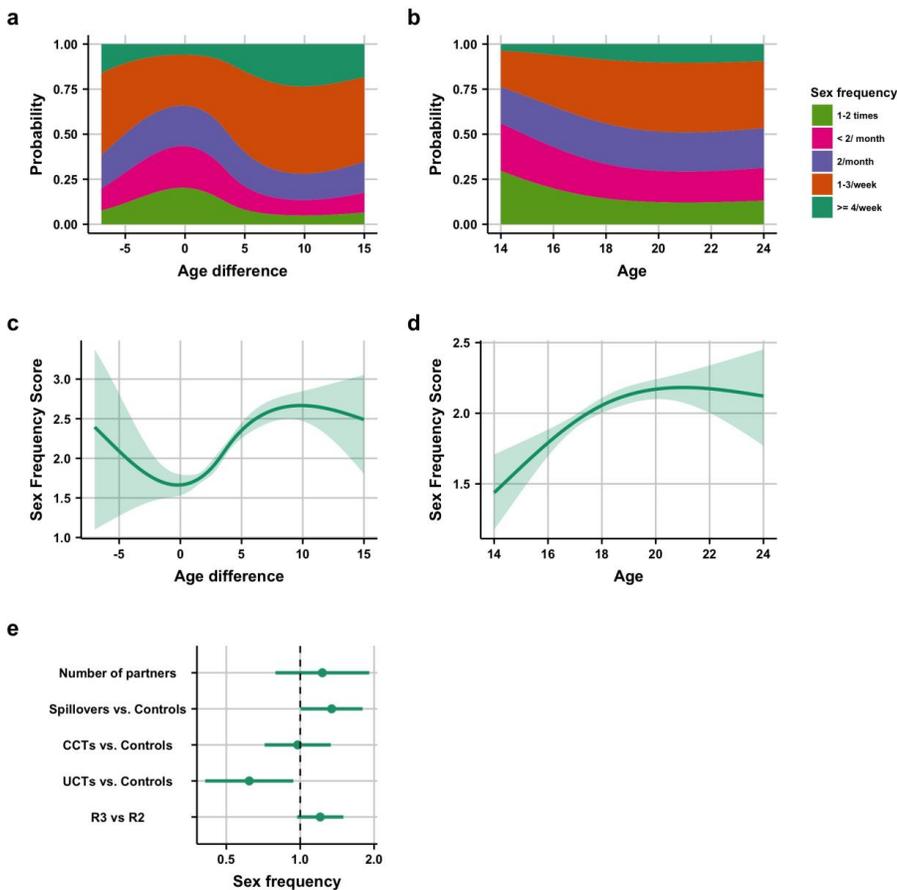
Figures 5a and 5c suggest a nonlinear relationship between age differences and increasing sex frequency, although at the extreme values of age difference the CIs are large (5c). When girls were of similar age as their partners sex frequency tended to be low and then increased rapidly with increasing age difference until the point where partners were around 7 or 8 years older, after which it stabilized. With regards to age (Figures 5b and 5d), when girls were young they tended to have a lower sex frequency. The probability that a girl had sex more frequently versus less frequently increased until around the age of 19, and thereafter remained relatively constant. Spillovers had higher odds of more frequent sex than controls (POR: 1.34; 95% CI: 1.00, 1.80), while UCTs had smaller odds of more frequent sex than controls (POR: 0.62; 95% CI: 0.41, 0.94). All girls had 1.21 times the odds (95% CI: 0.97, 1.50) of reporting more frequent sex after the intervention, compared to during the intervention.



**Figure 3. Results of linear mixed-effects model with age difference as the outcome.** Beta coefficient and 95% Confidence Interval (95% CI) plot for the (fixed) effects of age, study group, and round on age difference between a girl and her partner.



**Figure 4. Results of cumulative-link mixed model with condom use as the outcome.** In this model age difference was a spline term. a. cumulative probability of condom use categories for age difference. b. predicted effect of age difference on ordinal condom use score (scored as 0 for “never” up to 2 for “every time”), with shaded areas representing the 95% CI. c. proportional odds ratio (POR) and 95% CI plot for non-spline terms in the model.



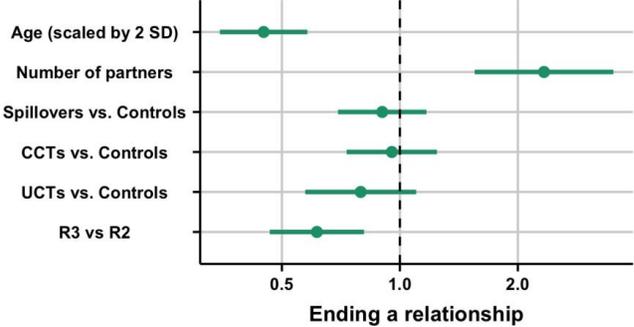
**Figure 5. Results of cumulative-link mixed model with sex frequency as the outcome.** Both age and age difference were spline terms in this model. a. cumulative probability of sex frequency categories for age difference. b. cumulative probability of sex frequency categories for age of participant. c. predicted effect of age difference on ordinal sex frequency score (scored as 0 for “1-2 times” up to 4 for “4 per week”) for age difference, with the shaded areas representing 95% CIs. d. ordinal sex frequency score for age of participant with the shaded areas representing 95% CIs. e. proportional odds ratio (POR) and 95% CI plot for non-spline terms in the model.

### *Relationship duration model*

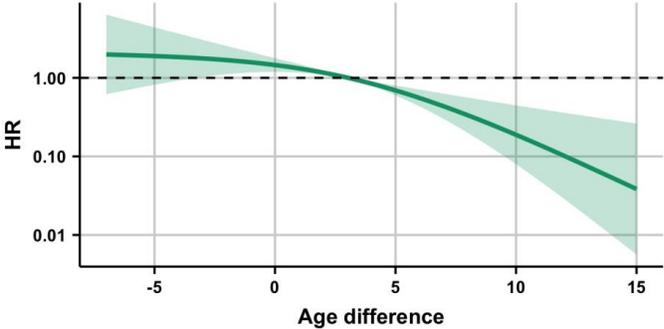
Figure 6a indicates that for every 4 years increase in age, the hazard of ending relationships decreased by 55% (HR: 0.45; 95% CI: 0.35, 0.58). The hazard of ending a relationship was 39% lower (HR: 0.61; 95% CI: 0.47, 0.81) for relationships reported in R3 compared to R2. There is no convincing evidence of an effect of the study group on the hazard of ending a relationship. Figure 6b shows that the hazard of ending a relationship was greater for girls more than 3 years older than their partners compared to girls who had the median age difference (3 years younger). This hazard gradually

decreased as the age differences became larger. The median relationship duration was shortest for girls one year older than their partners and longest for girls who were 10 years younger than their partners (Figure 6c).

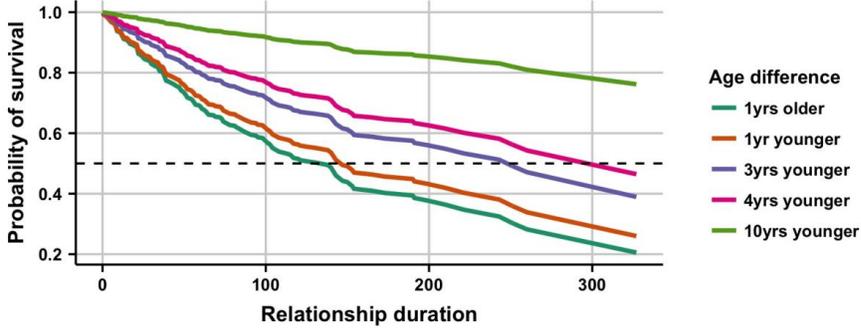
**a**



**b**



**c**



**Figure 6. Results of Cox proportional hazards model for relationship duration.** All of the plots represent the effects from the model, controlling for age, number of partners, study group and study round. In this model age difference was represented with a spline. a. coefficient plot of hazard ratios for ending relationships (HR and 95% CI) for all non-spline terms in the model. b. predicted HRs for age differences, with the median (age difference = 3) as the reference. c. expected survival curves for selected age differences (2.5th, 25th, 50th, 75th and 97.5th percentiles).

## Discussion

The SIHR study found that the prevalence of HIV and HSV-2 was lower among adolescent girls and young women in the intervention arms of a community randomised cash transfer trial (34). We found little support for our hypothesis that this result might be explained by differences in partner age between arms. Although we found that reported age differences in the combined intervention groups were on average smaller than in the control group during the intervention (R2), the effect was small and not statistically significant.

This contrasts with a study of government-administered social grants in South Africa, which found that the incidence of age-disparate relationships was 71% lower among girls who were part of households receiving grants (31). The differing findings between the study presented here and the one in South Africa might be partially explained by the differences in the way we operationalized age differences: while we used a continuous response, they dichotomized their response into those having or not having partners five or more years older than the participant. Moreover, there may have been different underlying motivations for entering age-disparate relationships in the two contexts. If schoolgirls in Zomba do not choose partners based upon financial reasons, then they would not have been incentivized to choose younger partners during the intervention, as we hypothesized.

In our study, large age differences in relationships were also associated with other behaviours potentially suggestive of heightened transmission risk, including lower levels of condom use, more frequent sex, and longer relationship durations. Our findings, along with the observation that the HIV prevalence among men in Malawi increases with increasing age up to 45 years old (16), support the hypothesis that age-disparate relationships provide more exposure to STIs, including HIV, for young women in Malawi.

While similar associations between age-disparate relationships and condom use (18–22) or sex frequency (20,21) have been noted in other studies, the relationship between age differences and duration of relationships has been less studied. We found that the older a man was compared to the girl, the less likely they were to terminate their relationship. Longer relationships may confer more risk of HIV transmission compared to shorter ones, even if condom use and sex frequency within the relationship were held constant, because there would be more time to transmit HIV. However, such effects could be negated if girls with shorter relationships had more partners in a given amount of time. Further study of the interaction between relationship duration, partner-turnover rate, and age differences in different age groups is needed to determine whether advantages of short duration in age-similar relationships are offset by choosing to have more sexual partners.

Interestingly, when the post-intervention period (R3) was compared to the intervention period there were larger age differences, less consistent condom use, more frequent sex,

and longer relationship durations. If the intervention was causing the young girls who received cash to participate in safer sexual behaviours, we would expect the riskier behaviours to increase when the girls stop receiving money. However, for the study groups not receiving money, we would expect the risky behaviours to remain constant during and after the cash transfer programs, because they did not have the monetary incentives to choose safer partners. We hypothesize that this could be explained by the presence of the Hawthorne Effect (49): because the controls and spillovers were aware of an ongoing study in their district they may have modified, or reported, their behaviour in a socially desirable way during the intervention. This effect has also been suggested to explain increased school attendance among the control communities of the South African HPTN 068 trial, measuring HIV incidence in young women in rural South Africa (50,51).

We also observed that older girls tended to have smaller partner age differences. Others have previously found that age differences between young women and their partners remain relatively constant as they age, usually varying between 4 and 7 years depending on the setting (18,24,52). While this is different from what we found, it demonstrates how age-mixing patterns are context-specific, and there may be different motivations and informal rules governing these choices from one setting to the next (53). We also noted that older girls tended towards less condom use, and were less likely to end relationships. Moreover, they had sex more frequently (although this trend plateaued in the young adults). One potential reason for our findings might be that as girls get older, men of similar age are more likely to have ways of earning money, thus making them more attractive partners. Also, as they enter more stable relationships we would expect growing trust within the relationship, thus leading to more regular sex and less consistent condom use (54).

This study has some limitations. The first is that there may be under-reporting of relationships due to social desirability bias. We observed that only 1108 out of 2907 schoolgirls reported a sexual relationship in either R2 or R3, with only 355 girls reporting a partner in both rounds. Face-to-face interviewing methods have been shown to result in under-reporting in sexual behaviour surveys (55–57). The effect of this bias on results, however, depends on whether unreported relationships differ systematically from those reported. Secondly, since most women did not report more than one relationship, we had limited ability to study partner turnover rate and multiple partner concurrency, both of which have been correlated with age-disparate relationships (7,13,17). Furthermore, some of the relationships reported in R2 and R3 may have been with the same partner, but the survey design did not allow for unique identification of relationships. We addressed this to the best of our ability by using random-effects models to account for potentially correlated relationships data reported by participants. Finally, there were some missing data on our key variable, age difference. Fortunately, only 8.5% of the relationships (127/1491) had missing data on this variable, and therefore we do not think the magnitude of the bias would be very large.

## Conclusions

Our analysis suggests that the primary mechanisms through which age-similar relationships prevent HIV in this population may be through increased condom use, lower sex frequency and a lower background HIV prevalence among male partners. Relationship duration may also play a role. Though we did not observe a conclusive study group effect on the age differences themselves, the increase in partner ages after

the programme ended compared to during the program is suggestive of potential Hawthorne effects. Thus, the HIV prevention benefits of the cash transfer intervention may have extended to those who were not receiving cash, since they also showed reduced prevalence of risky behaviours during the timing of the intervention. However, these effects, across all study groups, may be transient and only effective while cash transfers are taking place.

The quest for effective and cost-efficient interventions that prevent young women from acquiring HIV is ongoing. This analysis should encourage sustained action from policy makers in the health, education, and economic sectors to help create supportive socio-economic environments that facilitate safer relationship choices among young women while at the same time improving education levels and household income in communities where poor education, poverty and STI/HIV infection are pervasive challenges.

### **List of abbreviations**

HIV, Human Immunodeficiency Virus

STI, Sexually Transmitted Infection

HSV-2, Herpes Simplex Virus Type 2

SIHR, Schooling, Income, and Health Risk Study

EA, Enumeration Area

UCT, Unconditional Cash Transfer

CCT, Conditional Cash Transfer

R1, Round 1

R2, Round 2

R3, Round 3

CLMM, Cumulative Link Mixed Model

GAM, Generalised Additive Model

EDF, Effective Degrees of Freedom

DF, Degrees of Freedom

CI, Confidence Interval

$\beta$ ,  $\beta$ -coefficient

POR, Proportional Odds Ratio

HR, Hazard Ratio

## **Declarations**

### *Ethics approval*

Institutional Review Board approval to conduct this secondary analysis was obtained from the Stellenbosch University Health Research Ethics Committee (IRB0005239). The requirement for written consent was waived by the ethics committee because this was a secondary analysis, and no participants were contacted, nor did we have access to identifying information.

### *Consent for publication*

Not applicable

### *Availability of data and materials*

The datasets generated and/or analysed during the current study are available in The World Bank Central Microdata Catalogue repository, [http://microdata.worldbank.org/index.php/catalog/impact\\_evaluation](http://microdata.worldbank.org/index.php/catalog/impact_evaluation).

### *Competing Interests*

The authors declare that they have no competing interests.

### *Funding*

This work was supported by the Research Foundation Flanders – FWO [G0B4314N] and the International Clinics on Infectious Disease Dynamics and Data (ICI3D) Program under NIH award number R25GM102149 to JRC Pulliam and A Welte.

### *Author's Contributions*

RB conducted the statistical analysis, developed the figures and drafted the manuscript. WD and JD supervised and guided the analysis. All authors took equal part in the analysis design, interpretation of results, and revision of the manuscript.

### *Acknowledgements*

We would like to thank Sarah Baird at George Washington University, Berk Ozler at the World Bank, and Craig McIntosh at the University of California – San Diego for

access to the Schooling, Income, and HIV Risk Study data and their support in the data management and interpretation of preliminary results. Additionally, we thank Steve Bellan at University of Georgia and Brian Williams at SACEMA for reviewing drafts of the manuscript and providing feedback. Finally, we thank Michael Li from McMaster University for his contribution to writing functions used in the visualization of some of the models.

## References

1. Gouws E, Williams BG. Age-mixing and the incidence of HIV among young women. *Lancet HIV*. 2016;3018(16):10–1.
2. Delva W, Abdool Karim Q. The HIV epidemic in Southern Africa - Is an AIDS-free generation possible? *Curr HIV/AIDS Rep*. 2014;11(2):99–108.
3. Anderson RM, May RM, Ng TW, Rowley JT. Age-Dependent Choice of Sexual Partners and the Transmission Dynamics of HIV in Sub-Saharan Africa. *Philos Trans Biol Sci*. 1992;336(1277):135–55.
4. Garnett GP, Anderson RM. Sexually transmitted diseases and sexual behavior: insights from mathematical models. *J Infect Dis*. 1996;174 Suppl:S150-61.
5. Bershteyn A, Klein DJ, Eckhoff PA. Age-dependent partnering and the HIV transmission chain: a microsimulation analysis. *J R Soc Interface*. 2013;10(88):20130613.
6. d’Albis H, Augeraud-Véron E, Djemai E, Ducrot A. The dispersion of age differences between partners and the asymptotic dynamics of the HIV epidemic. *J Biol Dyn*. 2012;6(2):695–717.
7. 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*. 2002/06/12. 2002;359(9321):1896–903.
8. Pettifor AE, Rees H V, Kleinschmidt I, Steffenson AE, MacPhail C, Hlongwa-Madikizela L, et al. Young people’s sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. *AIDS*. 2005/09/02. 2005;19(14):1525–34.
9. Pettifor A, MacPhail C, Selin A, Gómez-Olivé FX, Rosenberg M, Wagner RG, et al. HPTN 068: A Randomized Control Trial of a Conditional Cash Transfer to Reduce HIV Infection in Young Women in South Africa---Study Design and Baseline Results. *AIDS Behav*. 2016;1–20.
10. Evans M, Risher K, Zungu N, Shisana O, Moyo S, Celentano DD, et al. Age-disparate sex and HIV risk for young women from 2002 to 2012 in South Africa. *J Int AIDS Soc*. 2016;19(1):1–16.
11. Harling G, Newell ML, Tanser F, Kawachi I, Subramanian S V, Barnighausen T. Do Age-Disparate Relationships Drive HIV Incidence in Young Women? Evidence from a Population Cohort in Rural KwaZulu-Natal, South Africa. *J Acquir*

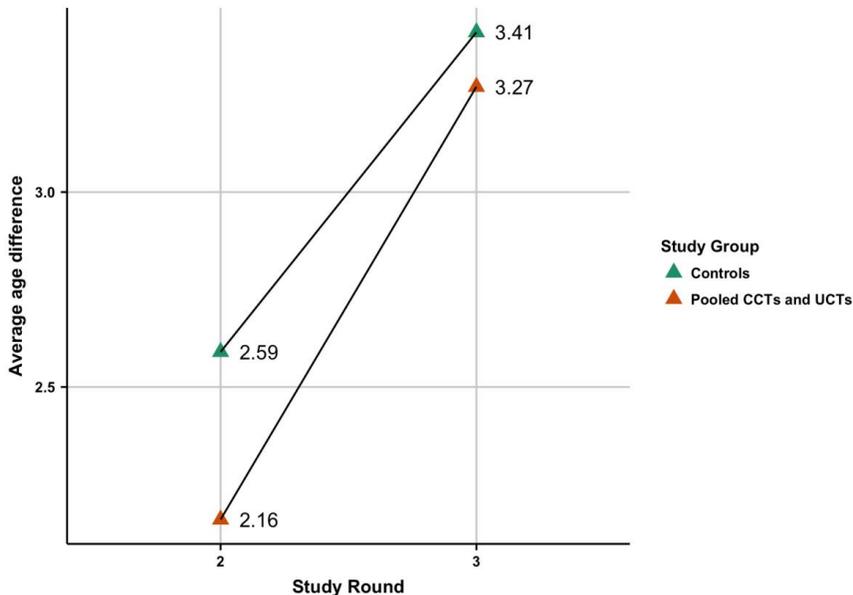
- Immune Defic Syndr. 2014;66(4):443–51.
12. Balkus JE, Nair G, Montgomery ET, Mishra A, Palanee-phillips T, Ramjee G, et al. Age-disparate partnerships and risk of HIV-1 acquisition among South African women participating in the VOICE trial. *J Acquir Immune Defic Syndr*. 2015;70(2):212–7.
  13. Street RA, Reddy T, Ramjee G. The generational effect on age disparate partnerships and the risk for human immunodeficiency virus and sexually transmitted infections acquisition. *Int J STD AIDS*. 2015;
  14. Gouws E, Stanecki KA, Lyerla R, Ghys PD. The epidemiology of HIV infection among young people aged 15-24 years in southern Africa. *Aids*. 2008;22(SUPPL. 4):S5–16.
  15. Shisana O, Rehle T, Simbayi L, Zuma K, Jooste S, Pillay-van-Wyk V, et al. South African national HIV prevalence, incidence, behaviour and communication survey, 2008: A turning tide among teenagers? Communication. Cape Town: HSRC Press; 2009.
  16. Office NS, Macro ICF. Malawi Demographic and Health Survey 2010. Zomba, Malawi Calverton, Maryland, USA: NSO and ICF Macro; 2011.
  17. Maughan-Brown B, Kenyon C, Lurie MN. Partner Age Differences and Concurrency in South Africa: Implications for HIV-Infection Risk Among Young Women. *AIDS Behav*. 2014;
  18. Beauclair R, Kassanjee R, Temmerman M, Welte A, Delva W. Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa. *Eur J Contracept Reprod Heal Care*. 2012/01/14. 2012;17(1):30–9.
  19. Maughan-brown B, Evans M, George G. Sexual Behaviour of Men and Women within Age-Disparate Partnerships in South Africa : Implications for Young Women ' s HIV Risk. *PLoS One*. 2016;1–16.
  20. Morrison-Beedy D, Xia Y, Passmore D. Sexual risk factors for partner age discordance in adolescent girls and their male partners. *J Clin Nurs*. 2013;22(23–24):3289–99.
  21. Ritchwood TD, Hughes JP, Jennings L, MacPhail C, Williamson B, Selin A, et al. Characteristics of Age-Discordant Partnerships Associated With HIV Risk Among Young South African Women (HPTN 068). *JAIDS J Acquir Immune Defic Syndr*. 2016 Aug;72(4):423–9.
  22. Toska E, Cluver LD, Boyes M, Pantelic M, Kuo C. From 'sugar daddies' to 'sugar babies': exploring a pathway between age-disparate sexual relationships, condom use, and adolescent pregnancy in South Africa. *Sex Health*. 2016;12(1):59–66.
  23. Nkosana J, Nkosana LBM. Intergenerational sexual relationships: The voices of men. *Integr J Br*. 2015;2(9):14–31.
  24. Beauclair R, Helleringer S, Hens N, Delva W. Age differences between sexual

- partners, behavioural and demographic correlates, and HIV infection on Likoma Island, Malawi. *Sci Rep.* 2016;6(May):36121.
25. de Oliveira T, Kharsany ABM, Gräf T, Cawood C, Khanyile D, Grobler A, et al. Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: a community-wide phylogenetic study. *Lancet HIV.* 2016;3018(16):1–10.
  26. Akullian A, Bershteyn A, Klein D, Vandormael A, Barnighausen T, Tanser F. Sexual partnership age-pairings and risk of HIV acquisition in rural South Africa. *AIDS.* 2017 Jun;31:1755–64.
  27. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med.* 2005;25(1):127–41.
  28. West RM. Generalised Additive Models. In: Tu YK, Greenwood D, editors. *Modern Methods for Epidemiology.* Dordrecht: Springer; 2012. p. 261–78.
  29. Irwin JR, McClelland GH. Negative Consequences of Dichotomizing Continuous Predictor Variables. *J Mark Res.* 2003;40(3):366–71.
  30. Altman DG, Royston P. The cost of dichotomising continuous variables. *Br Med J.* 2006;332(7549):1080.
  31. Cluver L, Boyes M, Orkin M, Pantelic M, Molwena T, Sherr L. Child-focused state cash transfers and adolescent risk of HIV infection in South Africa: a propensity-score-matched case-control study. *Lancet Glob Heal.* 2013;1(6):e362–70.
  32. Dunbar MS, Kang Dufour MS, Lambdin B, Mudekunya-Mahaka I, Nhamo D, Padian NS. The SHAZ! Project: Results from a Pilot Randomized Trial of a Structural Intervention to Prevent HIV among Adolescent Women in Zimbabwe. *PLoS One.* 2014;9(11):e113621.
  33. Hallfors DD, Cho H, Rusakaniko S, Mapfumo J, Iritani B, Zhang L, et al. The Impact of School Subsidies on HIV-Related Outcomes Among Adolescent Female Orphans. *J Adolesc Heal.* 2015;56(1):79–84.
  34. Baird SJ, Garfein RS, McIntosh CT, Ozler B. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. *Lancet.* 2012/02/22. 2012;379(9823):1320–9.
  35. Longfield K, Glick A, Waithaka M, Berman J. Relationships between older men and younger women: Implications for STIs/HIV in Kenya. *Stud Fam Plann.* 2004;35(2).
  36. Moore AM, Biddlecom AE, Zulu EM. Prevalence and meanings of exchange of money or gifts for sex in unmarried adolescent sexual relationships in sub-Saharan Africa. *Afr J Reprod Health.* 2008/05/07. 2007;11(3):44–61.
  37. Leclerc-Madlala S. Age-disparate and intergenerational sex in southern Africa: the dynamics of hypervulnerability. *AIDS.* 2008/12/17. 2008;22 Suppl 4:S17-25.
  38. Beauclair R, Delva W. Is younger really safer? A qualitative study of perceived risks and benefits of age-disparate relationships among women in Cape Town, South Africa. *PLoS One.* 2013;8(11).

39. Luke N. Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa. *Stud Fam Plann.* 2003/08/02. 2003;34(2):67–86.
40. Weiser SD, Leiter K, Bangsberg DR, Butler LM, Percy-de Korte F, Hlanze Z, et al. Food insufficiency is associated with high-risk sexual behavior among women in Botswana and Swaziland. *PLoS Med.* 2007/10/26. 2007;4(10):1589–97; discussion 1598.
41. Baird S, McIntosh C, Özler B. When the money runs out: Do cash transfers have sustained effects? Unpubl Manuscr. 2016;
42. Baird S, McIntosh C, Özler B. Cash or condition? Evidence from a cash transfer experiment. *Q J Econ.* 2011;126(4):1709–53.
43. Mpoola L., Chanza C, Nankuyu F, Kamela H, Kaudna L., Chirambo A, et al. Malawi: Zomba Urban Profile. 2011.
44. Baird S, Bohren A, McIntosh C, Ozler B. Designing Experiments to Measure Spillover Effects. Philadelphia, PA; 2014. Report No.: 14–032.
45. Baird S, Chirwa E, McIntosh C, Özler B. What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi, 3ie Impact Evaluation Report 27. New Delhi; 2015.
46. Solon G, Haider SJ, Wooldridge J. WHAT ARE WE WEIGHTING FOR? What Are We Weighting For? What Are We Weighting For? Cambridge, MA; 2013. (NBER Working paper series). Report No.: 18859.
47. Simon N. Wood. Generalized Additive Models: An Introduction with R. Taylor & Francis; 2006.
48. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
49. Parsons HM. What Happened at Hawthorne? *Science* (80- ). 1974;183(4128):922–32.
50. Rosenberg M. Selection and Hawthorne effects in HIV prevention trial among young South African women: Results from the HPTN 068 cohort. In: AIDS conference. Durban, South Africa; 2016.
51. Cluver L, Sherr L. Cash transfers -- magic bullet or fundamental ingredient? *Lancet Glob Heal.* 2016;4(12):e883–4.
52. Smith J, Nyamukapa C, Gregson S, Lewis J, Magutshwa S, Schumacher C, et al. The Distribution of Sex Acts and Condom Use within Partnerships in a Rural Sub-Saharan African Population. *PLoS One.* 2014;9(2):e88378.
53. Casterline JB, Williams L, Mcdonald P. Population Studies : A Journal of The Age Difference Between Spouses : Variations among Developing Countries The Age Difference Between Spouses " Variations among Developing Countries \*. (March 2013):37–41.
54. Marston C, King E. Factors that shape young people's sexual behaviour: a

- systematic review. *Lancet*. 2006;368(9547):1581–6.
55. Beauclair R, Meng F, Deprez N, Temmerman M, Welte A, Hens N, et al. Evaluating Audio Computer Assisted Self-Interviews in Urban South African Communities: Evidence for Good Suitability and Reduced Social Desirability Bias of a Cross-sectional Survey on Sexual Behaviour. *BMC Med Res Methodol*. 2013;13(11):11.
  56. Hewett PC, Mensch BS, Erulkar AS. Consistency in the reporting of sexual behaviour by adolescent girls in Kenya: a comparison of interviewing methods. *Sex Transm Infect*. 2004;80(suppl 2):ii43-ii48.
  57. Dolezal C, Marhefka SL, Santamaria EK, Leu CS, Brackis-Cott E, Mellins CA. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Arch Sex Behav*. 2011/05/24. 2012;41(2):401–10.

## Additional File 1



**Supplementary Figure 1. Interaction plot based upon Figure 3 from the manuscript.** This is an example interaction plot that demonstrates how to interpret the model effects for Figure 3 in the presence of interaction terms. For the example, we use the comparison between pooled CCTs and UCTs versus the Controls. The average age difference for Controls at R2 is the intercept (Average age difference = 2.59) of the model presented in Figure 3 of the main manuscript. The average age difference for Controls at R3 is the intercept plus the estimate for the time effect ( $2.59 + 0.82$ ). The average age difference for pooled CCTs and UCTs at R2 is the intercept plus the estimate for CCTs and UCTs vs Controls at R2 ( $2.59 - 0.43$ ). Finally, the average age difference for the pooled CCTs and UCTs at R3 is the addition of the intercept, the estimate for CCTs and UCTs vs Controls at R2, the estimate for the time effect, and the estimate for the interaction term ( $2.59 - 0.43 + 0.82 + 0.29$ ).

CHAPTER 7: THE ROLE OF AGE-MIXING PATTERNS IN  
HIV TRANSMISSION DYNAMICS: NOVEL  
HYPOTHESES FROM A FIELD STUDY IN CAPE TOWN  
SOUTH AFRICA

---



Photo credit: Dplanet: via Visualhunt.com / CC BY

## 7.1 AIMS OF ANALYSIS

---

Chapter 5 introduced methods for quantifying age-mixing patterns with observational data, and it additionally proposed a new individual-level indicator, called the bridge width, to describe the range of partner ages a person might have. We wanted to explore whether these new techniques for studying age-mixing at the population- and individual-level would be useful to our understanding of age differences and their relation to population HIV prevalence and individual HIV status (**Overall Objective 3**, See Section 1.6). Thus, in this final manuscript, we set out to describe the age-mixing pattern in the CTSBS population and determine if the patterns differed between HIV positive and HIV negative, men and women. Additionally, we explored whether being HIV positive was associated with participants having larger bridge widths in the year before the study. The manuscript that follows has been formatted for the *Epidemics* journal and is under review.

## **The Role of age-mixing patterns in HIV transmission dynamics: novel hypotheses from a field study in Cape Town, South Africa**

Roxanne Beauclair <sup>\*1,2</sup>, Niel Hens <sup>3,4</sup>, Wim Delva <sup>1,2,3,5</sup>

<sup>1</sup> International Centre for Reproductive Health, Ghent University, Gent, Belgium

<sup>2</sup> The South African Department of Science and Technology-National Research Foundation (DST-NRF) Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa

<sup>3</sup> Center for Statistics, I-BioStat, Hasselt University, Martelarenlaan 42, BE3500 Hasselt, Belgium

<sup>4</sup> Centre for Health Economics Research and Modelling Infectious Diseases and Centre for the Evaluation of Vaccination, Vaccine & Infectious Disease Institute, University of Antwerp, Antwerp, Belgium.

<sup>5</sup> Rega Institute for Medical Research, KU Leuven, Herestraat 49, 3000 Leuven, Belgium

\* Corresponding author: Roxanne Beauclair

+27 021 808 2589

roxanne.beauclair@gmail.com

## Abstract

**Background:** Age-disparate relationships are thought to put young women at increased risk of HIV, though current evidence is inconclusive. Studying population-level age-mixing patterns as well as individual-level measures of age difference variation may provide insight into the persistence and magnitude of the epidemic in South Africa.

**Methods:** We used data from a survey in Cape Town (n = 506) to describe age-mixing dynamics in the four population strata of HIV negative and HIV positive male and female participants. Mixed effects models were used to calculate the average increase in partner age for each year increase in age of participant, the average partner age for 15 year olds, and the between-subject and the within-subject standard deviation of partner ages. We conducted 2,000 bootstrap replications of the. Using negative binomial models, we also explored whether HIV status was associated with participants having a larger range in partner ages.

**Results:** HIV positive women had large variability in partner ages at the population level, and at the individual level had nearly three times the expected range of partner ages compared to HIV negative women. This pattern may increase the potential for HIV transmission across birth cohorts and may partially explain the persistence of the epidemic in South Africa. Young men, who have been previously absent from the age-disparity discourse, also choose older partners who may be putting them at increased risk of HIV infection.

**Keywords:** Age-disparate relationships, Age-mixing, HIV infection, South Africa, Sexual behaviour

### Highlights:

- Age-mixing compared among HIV positive and HIV negative men and women in Cape Town.
- At age 15, all population strata, including men, had older partners.
- Older female partners of young men have higher HIV prevalence than same age women.
- HIV positive women had large within-subject variability in partner ages.
- Large within-subject variability of partner ages facilitates HIV transmission between age cohorts.
- Young men with partners of older ages may also have increased risk of HIV infection.

## Introduction

Sexual relationships characterized by men being older than their female partners are thought to play a pivotal role in the magnitude and persistence of the HIV epidemic in Sub-Saharan Africa (Anderson et al., 1992; Chapman et al., 2010; Hallett et al., 2007). These relationships, commonly referred to as age-disparate relationships, may put young women at increased risk of acquiring HIV because the relationships also tend to be marked by power imbalances which may result in inconsistent condom use (Gregson et al., 2002; Maughan-brown et al., 2016; Morrison-Beedy et al., 2013; Ritchwood et al., 2016), higher sex frequencies per unit of time (Beauclair et al., 2016; Gregson et al., 2002; Morrison-Beedy et al., 2013; Ritchwood et al., 2016), longer relationship durations (Chapter 6), and multiple partnership concurrency (Gregson et al., 2002; Maughan-brown et al., 2016; Maughan-Brown et al., 2014; Ritchwood et al., 2016; Street et al., 2015). Additionally, in South Africa it has been noted that HIV prevalence for men peaks in older age groups compared to women (Shisana et al., 2014). This implies that if a younger woman chooses partners from an older age group, she will be more likely to select a partner that is HIV positive compared to if she was choosing a partner from a similar age (Gouws et al., 2008; Pettifor et al., 2005).

Despite great potential for age-disparate relationships to increase HIV transmission, some recent high-profile studies directly studying the link between age differences – in “primary” or “most recent” relationships – and HIV incidence did not find evidence of an association (Balkus et al., 2015; Harling et al., 2014; Street et al., 2015). Though, one study in Zimbabwe did find that increasing partner age differences were associated with increased hazards of acquiring HIV (Schaefer et al., 2017). Though, one study in Zimbabwe did find that increasing partner age differences were associated with increased hazards of acquiring HIV (Schaefer et al., 2017). Recent findings from Southern Africa seems to suggest that specific partner age pairings play an important role in HIV transmission (Akullian et al., 2017; de Oliveira et al., 2016), and all relationships (not only primary or most recent relationships) are important for continuing the cycle of HIV transmission in a population (Beauclair et al., 2016; de Oliveira et al., 2016). Specifically, the phylogenetic study by de Oliveira et al., found that young women under 25 years old most likely acquired HIV infection from men 25-40 years old, and then as they aged and entered the 25-40-year-old age cohort they infected age-similar partners with HIV.

Complicating this is the role that variation in partner age may play in HIV transmission at the population level. A modelling study that investigated how age-mixing patterns – population-level trends for how people choose sexual partners with regards to age – affect HIV transmission, found that increased variance in partner ages for the whole population led to larger values of the basic reproduction number of HIV (d’Albis et al., 2012). Once the variance was large enough, the mean age difference in the population no longer influenced the transmission dynamics. The role of partner age variability in HIV transmission dynamics has been understudied in real-world populations (Beauclair et al., 2016).

Thus far in age-mixing discourse most of the focus from observational studies has been on the individual-level risk posed by age differences in the most recent or primary relationship. Here, we are primarily interested in studying population-level attributes of age-mixing in real-world populations, and discussing how certain age-mixing patterns may create the conditions for HIV to be sustained in populations with high HIV

prevalence. Additionally, we hypothesize that individuals with a large range of partner ages may allow HIV to be acquired from one age group and transmitted to another age group. It is important to note that the study of age-mixing among men has been largely absent from the literature. To this purpose we conducted an analysis of both men and women using data from a sexual relationship survey in Cape Town, South Africa. We aimed to describe the age-mixing patterns in HIV negative and HIV positive male and female participants and to investigate whether being HIV positive was correlated with having a larger range of partner ages during the year prior to the survey.

## **Methods**

### *Data*

The data for this analysis come from a cross-sectional sexual behaviour survey we administered from June 2011 to February 2012 in three urban communities of Cape Town, South Africa that had a high HIV prevalence. The three study communities are designated “townships”: locations where “black” South Africans were allowed to live during Apartheid. These townships were characterized by high unemployment, low rates of graduation from secondary schools, as well as densely populated formal houses and shack dwellings (Department, 2012; Muzondo et al., 2004).

The survey employed Audio Computer-Assisted Self-Interviewing (ACASI) and collected in-depth sexual relationship histories of the participants for the year prior to the survey. Participants could report up to 15 partners.

The survey participants were randomly sampled from a roster of participants in the Zambia South Africa TB and AIDS Reduction (ZAMSTAR) prevalence survey, which took place in 2010, approximately one year before the sexual behaviour survey. In the ZAMSTAR prevalence survey the HIV status of participants was obtained via Abbott Determine HIV-1/2 screening tests (Ayles et al., 2008). These HIV test results were subsequently linked via a unique, anonymous ID to their sexual behaviour survey results. Thus, most of the participants that opted to take an HIV test would have known their HIV test results prior to taking the sexual behaviour survey. Knowledge of their HIV status may have influenced how they reported information about partnerships during the survey. Full details of the Cape Town sexual behaviour survey study design, sampling procedures, as well as, contact and response rates of have been described elsewhere (Beauclair et al., 2015, 2013, Delva et al., 2013, 2011).

All together 878 participants took the sexual behaviour survey. We wanted to focus this analysis on the black, heterosexual and sexually active population, since this is one of the demographic groups most affected by HIV in South Africa. Therefore, we excluded participants whose stated sexual preferences were for the same or both genders ( $n = 56$ ), those who did not identify as black ( $n = 206$ ), and those not reporting any sexual partners in the previous year ( $n = 110$ ). This left us with 506 participants reporting on 870 relationships.

### *Statistical Analysis*

Our data had missing observations for key participant-level variables: age ( $n = 17$ , 3.36%), sex ( $n = 13$ , 2.57%), and HIV status ( $n = 90$ , 17.79%). For partner age 7.47% of relationships ( $n = 65$ ) had missing values. A discussion of assumptions regarding the missingness mechanisms, as well as, comparisons of participant characteristics for key variables by missingness status can be found in the Supplementary Web Appendix 1. In order to account for the selection bias and loss of efficiency that otherwise would have been introduced by the missingness (Rubin, 1976), we imputed 100 datasets using multivariate imputation by chained equations with a random-forest algorithm (MICE-RF) (van Buuren et al., 2011; White et al., 2011). The random forest algorithm has been demonstrated to produce more accurate imputations when non-linear associations exist among variables used in analytic models (Penone et al., 2014; Stekhoven and Bühlmann, 2012). The “mice” and “CALIBERrfimpute” packages in R were used for MICE-RF (Shah et al., 2014; van Buuren et al. 2011). A list of the variables we included in our imputation model can be found in Supplementary Web Appendix 2.

We wanted to describe the age-mixing pattern in four strata of participants: HIV positive men (HPM), HIV positive women (HPW), HIV negative men (HNM), and HIV negative women (HNW). To do this, we used linear mixed-effects models to regress the age of the partner on the age of the participant for each dataset. These models also accounted for the hierarchical nature of the data by containing a random intercept for the participant (Bates, 2010; Bolker et al., 2009; Gelman, 2007).

The average increase in partner age for each year increase in age of participant (model slope or  $\beta$ -coefficient), the average partner age for 15-year-olds (population-level intercept), the between-subject variability (between-subject standard deviation, BSSD), and the within-subject variability (residual, i.e. within-subject standard deviation, WSSD) were computed for each stratum using these models. We centred our model on 15 year olds because they represent the youngest age in the dataset and we wanted to understand how individuals who are only starting to become sexually active choose partners with respect to age. We aimed to compare these features of the age-mixing pattern for HNM vs. HPM, HNW vs. HPW, HNM vs HNW, and HPM vs HPW.

Since some features lack a standard calculation for the 95% confidence intervals (95% CIs) (e.g. MI-averaged WSSDs, or the differences between two WSSDs), we created 2,000 participant-level bootstrap replicates for each imputed dataset ( $m = 100$ ), resulting in a total of 200,000 datasets (Davison and Hinkley, 1997; Efron, 1987). We applied the linear mixed-effects model for each population stratum of the 200,000 datasets and captured the age-mixing features. Then, we calculated the differences between those model features for the comparisons we were interested in and subsequently created 95% CIs using the percentile method.

Among the participants who reported more than one partner, we examined whether participant HIV status was associated with a large range of partner ages. We created an indicator of this range, we call a “bridge width”. The bridge width is defined as the number of years difference between the maximum and the minimum partner age of the participant (Beauclair et al., 2016). First, we plotted the distribution of bridge widths. Then we treated bridge width as an over-dispersed count variable and regressed it on HIV status using negative binomial regression. We stratified the models by sex and adjusted for the age of the participants. We hypothesized that age was related to bridge width in a nonlinear way, and thus, we used a generalized additive model with a smooth term for age (Hastie and Tibshirani, 1990; Wood, 2006). The exponentiated model

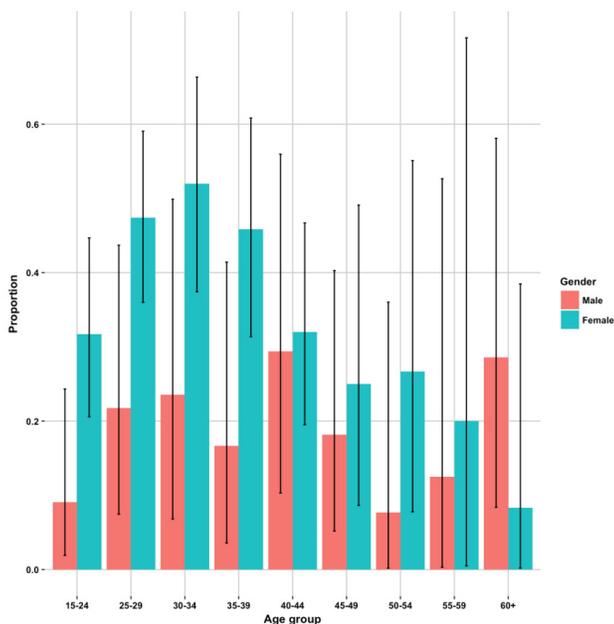
coefficient for HIV status is referred to here as the expected bridge width ratio (EBWR). EBWRs and 95% CIs from the imputed datasets were pooled according to Rubin's Rules (Rubin, 1987).

All analyses were conducted in R statistical software (R Development Core Team, 2014). In this paper, we present the results from the analysis with imputed data. However, relationship characteristics from the original dataset, as well as the complete case analyses can be found in Supplementary Web Appendix 3.

## Results

Of the relationships that did not have missing HIV status or gender data, 32.6% belonged to HNW ( $n = 219/671$ ), 35.7% to HNM ( $n = 240/671$ ), 24.7% to HPW ( $n = 166/671$ ), and 6.7% to HPM ( $n = 46/671$ ). Nearly 23% of relationships had either missing gender or HIV status (199/870).

We randomly selected one of the 100 imputed data sets to visualize the HIV prevalence in our Cape Town study population (Figure 1). In most age groups, women had a higher HIV prevalence than men. For women aged 15-24, HIV prevalence was nearly 3.5 times that of the men in that same age group (31.7% vs. 9.1%). This disparity was striking in 25-29-year-olds, with 47.4% of women infected and only 21.7% of men infected. HIV prevalence peaked in women 30-34 years old (52.0%) and in men who were 40-44 years old (29.4%).



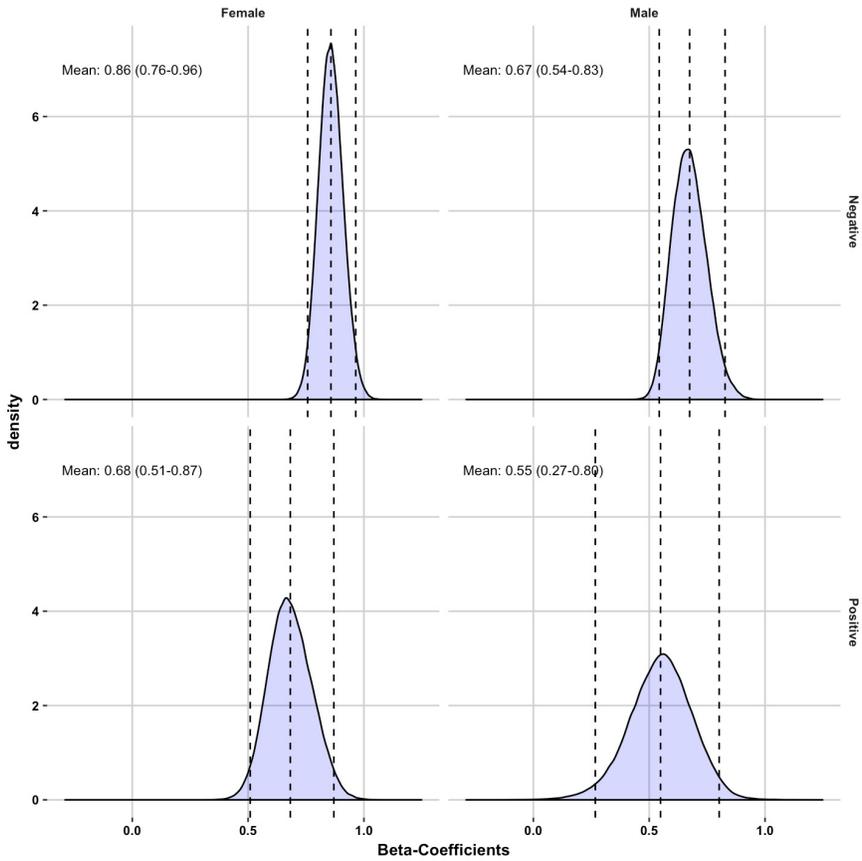
**Figure 1.** HIV prevalence by sex and age in randomly selected imputed dataset with 95% exact confidence intervals.

Figure 2 shows the distributions of model  $\beta$ -coefficients, which represent how fast the partner ages grow as participants' ages increase. For all population strata, the average partner ages increased with each year increase in participant age. HPM had the smallest average model slope: for each year increase in age of participants the partner age was only expected to increase by 0.55 years (95% CI: 0.27 - 0.8). The HNW had the largest model slopes and thus their partner ages were expected to become 0.86 years larger for each year increase in participant age (95% CI: 0.76 - 0.96) (Figure 2). Most of the distributions seem to overlap to a great extent and in Table 1 we see there were only moderate differences in  $\beta$ -coefficients between the groups, with all but one of the 95% CIs around these differences overlapping with zero. HNW had steeper slopes than HNM (Difference: 0.18; 95% CI: 0 - 0.35). Also in Table 1 we see that for the HNM and HNW the model intercept difference was approximately 4 years (95% CI: 0.78 - 8.36), with 15-year-old men and women in those groups having partners who were 19 (95% CI: 15 - 21) and 23-years-old (95% CI: 21 - 25) on average, respectively (Figure 3). If we take the mean model slopes and intercepts for each subpopulation the average partner age at 25 years old is predicted to be 32 years for HNW, 25 years for HNM, 32 years for HPW, and 28 years for HPM. The predicted partner ages for 40 year olds are: 44 years for HNW, 35 years for HNM, 42 years for HPW, and 37 years for HPM.

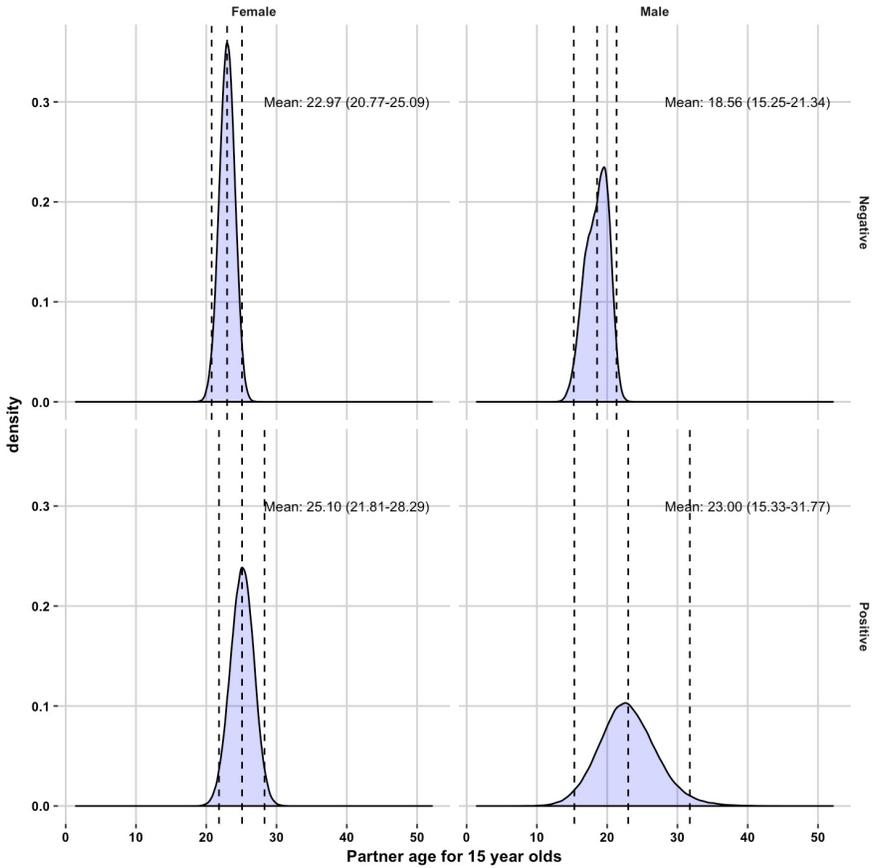
**Table 1: Differences between the age-mixing pattern features, derived from the linear mixed-effects models.** We calculated the differences between model features for HIV negative men (HNM) and HIV positive men (HPM), HIV negative women (HNW) and HIV positive women (HPW), HNM and HNW, as well as, HPM and HPW, for 100 imputed datasets, each bootstrapped 2000 times. We calculated 95% confidence intervals (CIs) using the percentile method.

Comparisons	$\beta$ -coefficient	Intercept	BSSD	WSSD
HNM vs. HNW	0.18 (0.00 — 0.35)	4.40 (0.78 — 8.36)	1.05 (-2.82 — 4.21)	1.01 (-1.11 — 3.28)
HPM vs. HPW	0.13 (-0.18 — 0.47)	2.10 (-7.23 — 10.50)	3.51 (-3.29 — 9.25)	1.47 (-3.62 — 6.17)
HPM vs. HNM	0.13 (-0.17 — 0.44)	4.43 (-4.01 — 13.85)	2.2 (-4.09 — 9.20)	2.03 (-2.02 — 6.54)
HPW vs. HNW	0.18 (-0.04 — 0.38)	2.13 (-1.87 — 6.10)	4.66 (1.58 — 7.86)	4.51 (0.90 — 7.79)

HNM, HIV Negative Men  
 HNW, HIV Negative Women  
 HPM, HIV Positive Men  
 HPW, HIV Positive Women  
 BSSD, Between-subject Standard Deviation  
 WSSD, Within-subject Standard Deviation



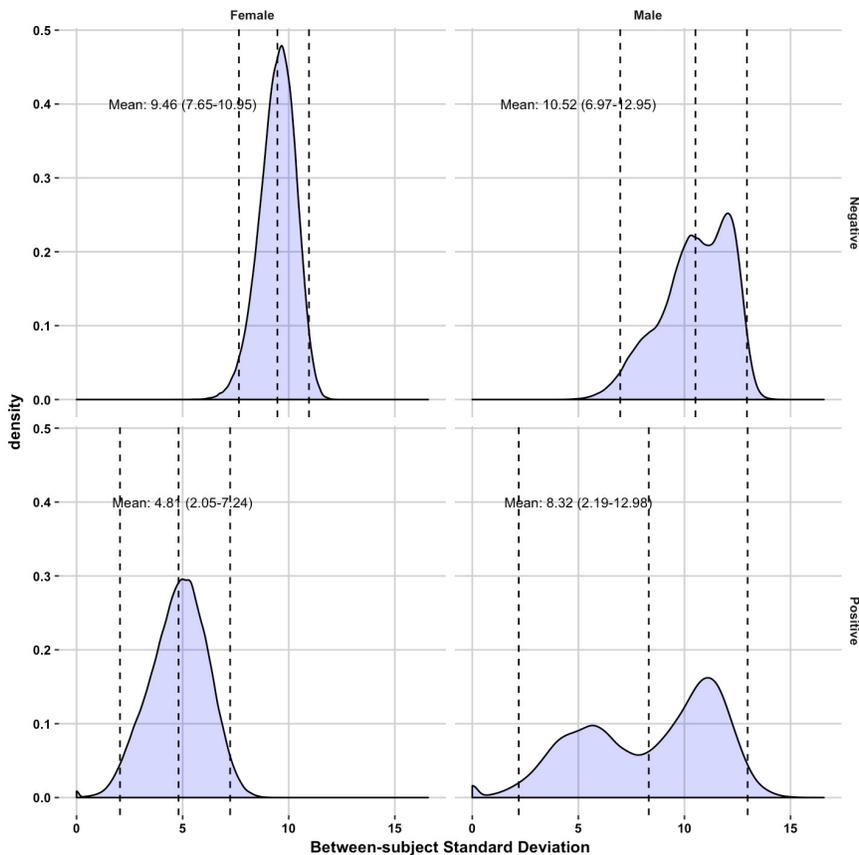
**Figure 2. Distribution of model slopes, derived from the linear mixed-effects models.** The models were applied to 100 imputed datasets, each bootstrapped 2000 times. The vertical dashed lines represent confidence intervals obtained through the percentile method. The mean of all the model slopes (beta-coefficients) represents the average increase in partner age for each year increase in participant age for that stratum.



**Figure 3. Distribution of model intercepts, derived from the linear mixed-effects models.** The models were applied to 100 imputed datasets, each bootstrapped 2000 times. The vertical dashed lines represent confidence intervals obtained through the percentile method. The mean of all the model intercepts represents the average predicted partner age for 15-year-olds in that stratum.

HNM had the largest BSSD (BSSD: 10.52; 95% CI: 6.97 - 12.95) while HPW had the lowest (BSSD: 4.81; 95% CI: 2.05 - 7.24) (Figure 4). There appears to be a wide and bimodal distribution of BSSDs in HPM. In the complete case data, there were only 46 relationships in this subpopulation, with the minimum and maximum partner ages being 22 and 79, respectively. We believe those extreme values were oversampled in some bootstrap replicates because of the small sample size, thus resulting in clusters of results with both small and large BSSDs. Table 1 shows that there was a relatively

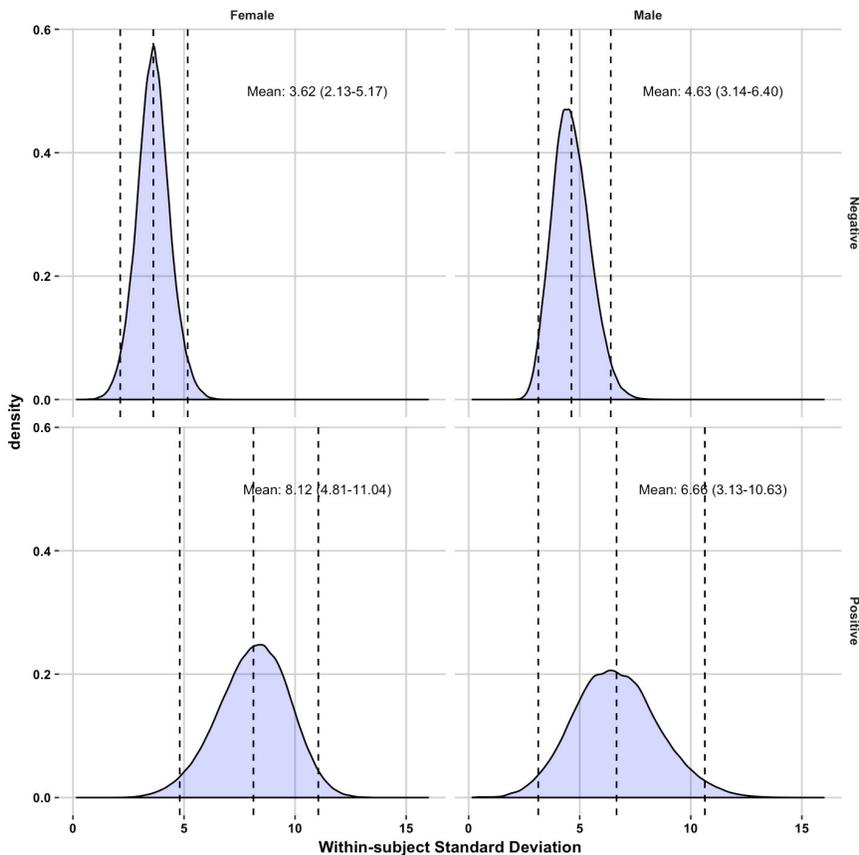
large and significant difference between the BSSD in HPW and HNW (Difference: 4.66; 95% CI: 1.58 - 7.86).



**Figure 4. Distribution of between-subject standard deviations (BSSD) of partner ages, derived from the linear mixed-effects models.** The models were applied to 100 imputed datasets, each bootstrapped 2000 times. The vertical dashed lines represent confidence intervals obtained through the percentile method. The mean BSSD represents the average variability in partner age that can be seen among individuals in that stratum.

Figure 5 shows that the WSSD was the largest for HPW (WSSD: 8.12; 95% CI: 4.81 - 11.04) and the smallest for HNW (WSSD: 3.62; 95% CI: 2.13 - 5.17), with the difference between the two groups being 4.51 (95% CI: 0.90 - 7.79) (Table 1). For the other group

comparisons, the WSSDs appear to be similar and their distributions have a large degree of overlap.



**Figure 5. Distribution of within-subject standard deviations (WSSD) of partner ages, derived from the linear mixed-effects models.** The models were applied to 100 imputed datasets, each bootstrapped 2000 times. The vertical dashed lines represent confidence intervals obtained through the percentile method. The mean WSSD represents the average within-subject variability of partner age for individuals within each stratum.

In the complete study population 155 participants reported having had more than one relationship in the previous year (30.63%). Figure 6 shows the distribution of bridge widths – range in partner ages – by population strata. In all subpopulations, the distributions were right-skewed, with most of the bridge widths in all imputations

spanning less than 20 years. In both HIV positive groups, there appears to be fatter tails, implying that those with HIV may have larger bridge widths, although there is more uncertainty in the imputations of HPM. The results of our individual-level negative-binomial models are presented in Figures 7 and 8. Across all imputations, having an HIV positive status increased the expected bridge width of women by almost three times that of those who were HIV negative (EBWR: 2.76; 95% CI: 1.16 - 6.59). Among men there also seems to be an increased expectation of larger bridge widths in HIV positive participants compared to HIV negative participants, although when all of the estimates were pooled for the imputed data sets, the evidence did not appear to be as convincing as the evidence for women (EBWR: 1.93; 95% CI: 0.69 - 5.41).

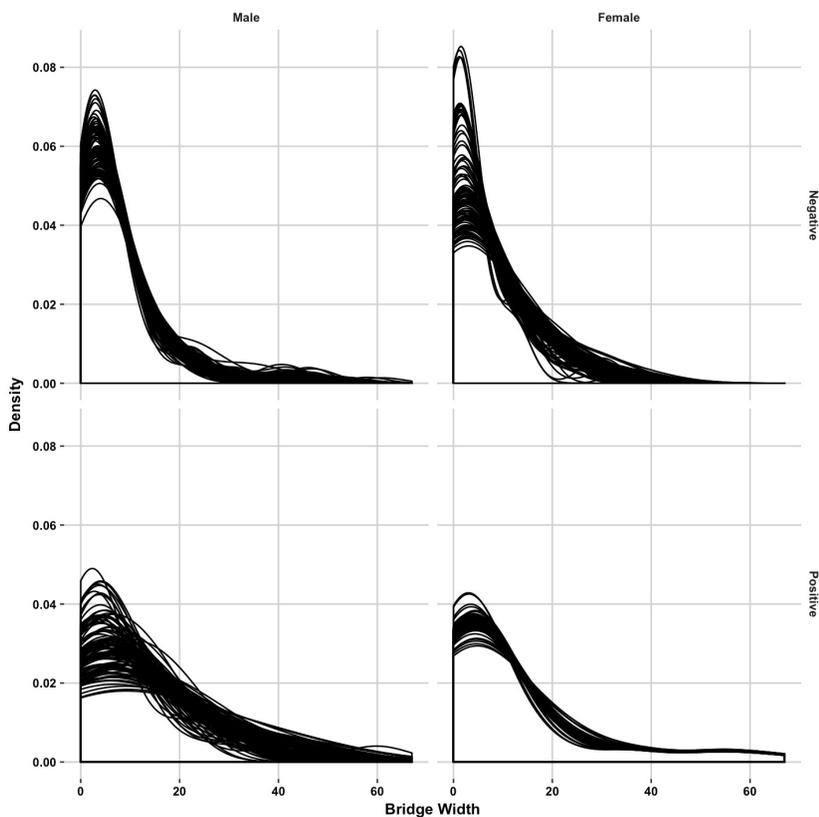
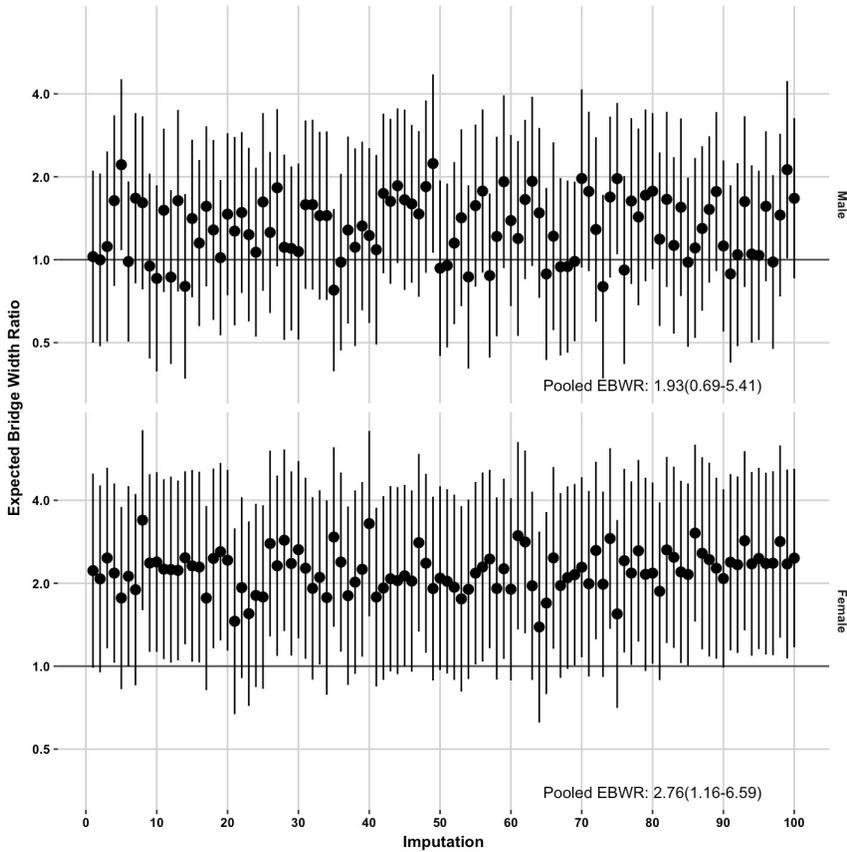


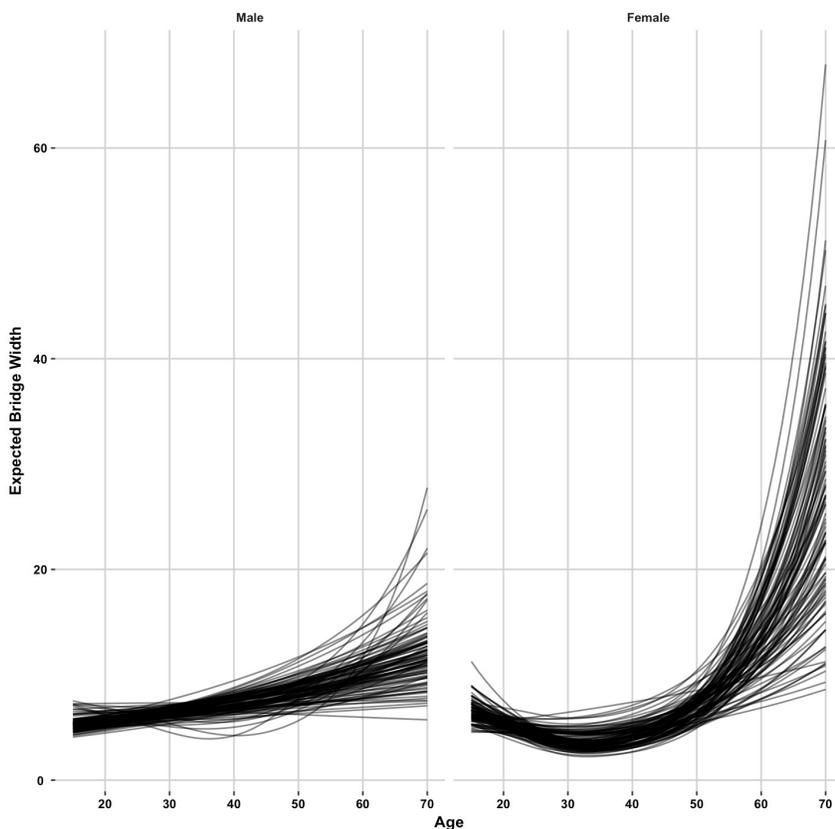
Figure 6. Un-adjusted distribution of bridge widths among those reporting more than one partner for each imputed dataset, by sex and HIV status.



**Figure 7. Model coefficients for relationship between HIV and bridge width, for each of the 100 imputations.** The Expected Bridge Width Ratio (EBWR) is the ratio of expected bridge widths for those who are HIV positive compared to those who are HIV negative. The EBWR is derived from the negative binomial model that also adjusts for the age of the participant. The EBWR that is annotated within the plots was pooled across imputations according to Rubin’s Rules. The individual EBWRs that are plotted for each imputation, demonstrate the amount of variability seen in the imputations.

Finally, in both men and women there appears to be a relationship between increasing participant age and larger bridge widths, although the uncertainty in the imputations is reflected by there being large variation in expected bridge widths at older ages. Indeed, in one randomly selected imputed dataset, there were only 22 participants (4.3%) with ages between 60 and 70 years. For men, the increase in the expected partner age ranges were not as large as for women. For women, in most of the imputed data sets, the relationship between age and expected bridge widths was nonlinear, with the youngest

and oldest participants having larger expected bridge widths, and those between 25 and 45 years old having smaller bridge widths.



**Figure 8. Expected bridge widths for different values of age (smooth term), by gender.** The expected bridge widths are predicted from the negative binomial model that also adjusts for the HIV status of participants. There is a line for each of the 100 models resulting from the imputations.

## Discussion

Our study of heterosexual, black participants in urban Cape Town communities with high HIV prevalence sheds light on key features of the age-mixing patterns in HIV positive and negative men and women. In particular, there were some important findings with implications for the role women play in HIV transmission.

Perhaps the most novel insight from this study was that the within-subject variability of partner ages among all HPW was the largest out of all four population strata. Moreover, there was a detectable difference between their within-subject variability and that of the HNW. This finding was also reinforced by our individual-level finding that HPW had nearly 3 times the expected range in partner ages compared to HNW. The HPM also had approximately two times the expected bridge width of HNM, though these results were not statistically significant. It is important to note that a large within-subject variability in partner ages does not necessarily mean that the women always have to have partners older than themselves. It is also possible that the HPW in our sample acquired HIV from age-similar partners. The large within-subject variability in partner ages in this group means that these women may have great potential to transmit HIV to men in other age cohorts, thereby allowing HIV to be sustained in a population (d'Albis et al., 2012). Interestingly, we witnessed the opposite effect for between-subject variability: the HPW had the smallest, and this was significantly different from the HNW. A larger fraction of HPW in our sample reported more than one partner compared to HNW. This may have resulted in the model being unable to distinguish which part of the overall variation in partner ages was due to the within component versus between component among HNW. Thus, the model might have attributed too much variation to differences between participants among HNW. We believe more research is needed to explore the utility of BSSDs for understanding HIV transmission dynamics.

Another important finding is that for all population strata, 15-year-olds chose older partners, on average. However, in the HIV negative population women chose significantly older partners than men did (23 years vs. 19 years). Critically, this demonstrates that in populations of young people, women are not the only ones choosing older partners. We also noted that in this study population the HIV prevalence among young women was many times greater than the HIV prevalence in age-similar men, indicating that if both young men and young women are choosing partners from their own age group the men will be more likely to choose an HIV positive partner, than the women. Most recent epidemiological studies examining the associations between age differences and HIV incidence have failed to include young men in their analyses, choosing to focus only on women in age-disparate relationships (Balkus et al., 2015; Harling et al., 2015, 2014; Schaefer et al., 2017; Street et al., 2015). However, there are a few studies that suggest that men in relationships with older or age-similar women may have increased odds of having HIV or a curable STI (Beauclair et al., 2016; Gregson et al., 2002; O'Leary et al., 2015). We also note that men in urban communities are highly mobile, and they may be choosing female partners outside of the study communities. The women from those communities may not have the same HIV prevalence profiles as the women in our study community.

Among HIV-negative participants, there was a marginally significant difference between men and women for the effect of age on the average partner age choice. Indeed, we saw that young men in this group were younger than their female partners when they were 15, but by the time they were 40 years old they were on average 5 years older than their female partners, while women remained younger than their male partners at 40 years. This finding that as men get older, their age differences grow (older than the woman) is similar to other findings (Beauclair et al., 2016, 2012; Smith et al., 2014). While in other populations, women have been observed to have constant age differences as age increases (Beauclair et al., 2016, 2012; Becker et al., 2014), in this study we observed, similar to Ott et al., that as women grew older, their age differences decreased

(Ott et al., 2011). We did not detect significant differences between the HIV negative and the HIV positive participants of either gender, and thus we hypothesize that this particular feature of the age-mixing pattern may not play as important a role in maintaining endemicity of HIV.

The results we present here are congruous with the findings from de Oliveira et al. who found in their phylogenetic analysis from KwaZulu-Natal, South Africa that HIV transmission pathways can be complex and involve cycles of transmission between different genders and age groups (de Oliveira et al., 2016). While our study was limited because we could not determine who infected whom, we have evidence that suggests young women below the age of 25, and older women above the age of 45, who have large bridge widths may be playing key roles in transmission. There have been many qualitative investigations about why young women are motivated to engage in age-disparate relationships (Beauclair and Delva, 2013; Hawkins et al., 2012; Leclerc-Madlala, 2008; Luke, 2003; Moore et al., 2007; Weiser et al., 2007). However, there is far less understanding about why older women have relationships with younger men, and vice versa (Kuate-Defo, 2008; Phaswana-Mafuya et al., 2014).

A key limitation of our study was that the sexual behaviour survey may have resulted in social desirability bias, through underreporting or over-reporting of sexual partners, in the cases of women and men, respectively. Additionally, we hypothesize that women with an HIV diagnosis might have been more likely to report age-similar partner ages, or refrain from answering questions about partner age. In this context, older men have a reputation for preferring to not use condoms during sex (Beauclair and Delva, 2013), and thus women who were HIV positive may have been more apprehensive towards reporting the ages of these partners, especially if they thought they may have contracted HIV from that partner. Fortunately, we were able to use techniques in the design of the survey, and in our analysis to counteract these forms of bias. We used ACASI methods to conduct the survey, which have been shown in other studies to reduce social desirability bias in different African settings (Dolezal et al., 2012; Hewett et al., 2004; Kissinger et al., 1999). Moreover, in a previously published paper using this data, we demonstrated how the ACASI interviewing modality likely resulted in more reporting of sexual partners and sensitive behaviours in comparison to other sexual behaviour surveys using face-to-face-interviewing (Beauclair et al., 2013). Finally, we also corrected for the bias that might have resulted from missing partner age data among HIV-positive participants, by using MICE-RF statistical methods.

There is also a possibility that the HIV status variable may have had missing observations because the participants were HIV positive. In this scenario, MICE probably did not completely correct for that form of selection bias, and thus, a sensitivity analysis may be warranted. More research is necessary on how to conduct sensitivity analyses in the presence of MICE-RF imputations.

Additionally, we were restricted by the small total sample size, particularly of HIV positive participants. This may have resulted in a limited ability to detect differences between the HIV positive participants and other population strata. Unfortunately, it is often difficult to find large sexual behaviour datasets where participants have felt comfortable enough to report multiple partners in order to study these dynamics. A potential way forward is to conduct simulation studies, which do not have the sample size constraints of real-world data. Ultimately, it is important to note that these findings may only be generalizable to other urban, disadvantaged communities within

Southern Africa that have similar HIV prevalence profiles among men and women of different age groups.

## Conclusions

While age-mixing has garnered a lot of recent attention for the role it may play in HIV transmission, most of the studies have focused on the individual-level perspective of young women in age-disparate relationships. To our knowledge, this exploratory study is the first to try describe several real-world population-level features of age-mixing patterns in HIV negative and positive populations, and to see if they differ in systematic ways. These differences, or lack thereof, furnish insight into the complex transmission dynamics that occur between genders at different ages, and could possibly lead to sustained HIV epidemics. The hypotheses we generated about which features of age-mixing patterns are important for HIV epidemics, like within-subject variability, can be confirmed through simulation studies, calibrated to real-world populations such as the ones described here. These simulations could then be used to evaluate how different age-mixing pattern characteristics affect the individual-level and population-level dynamics of HIV transmission over short and longer time scales.

## Acknowledgements

We would like to thank the fieldworkers who helped to carry out this study. We also acknowledge assistance from Nulda Beyers and Rory Dunbar from the Desmond Tutu TB Foundation at Stellenbosch University. They were instrumental with helping to design the study, providing us with a sampling frame and subsequently linking the ZAMSTAR HIV results to our sexual behaviour survey.

## Funding Sources

Funding: This work was supported by the Research Foundation - Flanders (FWO) [G091210N; G0B4314N; W002514N] and the Flemish Interuniversity Council (VLIR) [ZEIN2010PR375].

## References

- Akullian, A., Bershteyn, A., Klein, D., Vandormael, A., Barnighausen, T., Tanser, F., 2017. Sexual partnership age-pairings and risk of HIV acquisition in rural South Africa. *AIDS* 31, 1755–1764. doi:10.1097/QAD.0000000000001553
- Anderson, R.M., May, R.M., Ng, T.W., Rowley, J.T., 1992. Age-Dependent Choice of Sexual Partners and the Transmission Dynamics of HIV in Sub-Saharan Africa. *Philos. Trans. Biol. Sci.* 336, 135–155.
- Ayles, H.M., Sismanidis, C., Beyers, N., Hayes, R.J., Godfrey-Faussett, P., 2008. ZAMSTAR, The Zambia South Africa TB and HIV Reduction Study: design of a 2

- x 2 factorial community randomized trial. *Trials* 9, 63. doi:10.1186/1745-6215-9-63
- Balkus, J.E., Nair, G., Montgomery, E.T., Mishra, A., Palanee-phillips, T., Ramjee, G., Panchia, R., Selepe, P., Richardson, B.A., Chirenje, Z.M., Marrazzo, J.M., 2015. Age-disparate partnerships and risk of HIV-1 acquisition among South African women participating in the VOICE trial. *J. Acquir. Immune Defic. Syndr.* 70, 212–217. doi:10.1097/QAI.0000000000000715
- Bates, D.M., 2010. *lme4: Mixed-effects modelling with R*. Springer.
- Beauchair, R., Delva, W., 2013. Is younger really safer? A qualitative study of perceived risks and benefits of age-disparate relationships among women in Cape Town, South Africa. *PLoS One* 8.
- Beauchair, R., Hellingranger, S., Hens, N., Delva, W., 2016. Age differences between sexual partners, behavioural and demographic correlates, and HIV infection on Likoma Island, Malawi. *Sci. Rep.* 6, 36121. doi:10.1038/srep36121
- Beauchair, R., Hens, N., Delva, W., 2015. Concurrent partnerships in Cape Town, South Africa: race and sex differences in prevalence and duration of overlap. *J Int AIDS Soc* 18, 19372. doi:10.7448/IAS.18.1.19372
- Beauchair, R., Kassanje, R., Temmerman, M., Welte, A., Delva, W., 2012. Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa. *Eur J Contracept Reprod Heal. Care* 17, 30–39. doi:10.3109/13625187.2011.644841
- Beauchair, R., Meng, F., Deprez, N., Temmerman, M., Welte, A., Hens, N., Delva, W., 2013. Evaluating Audio Computer Assisted Self-Interviews in Urban South African Communities: Evidence for Good Suitability and Reduced Social Desirability Bias of a Cross-sectional Survey on Sexual Behaviour. *BMC Med Res Methodol* 13, 11.
- Becker, M., Vignoles, V.L., Owe, E., Easterbrook, M.J., Brown, R., Smith, P.B., Bond, M.H., Regalia, C., Manzi, C., Brambilla, M., Aldhafri, S., Gonzalez, R., Carrasco, D., Paz Cadena, M., Lay, S., Schweiger Gallo, I., Torres, A., Camino, L., Ozgen, E., Guner, U.E., Yamakoglu, N., Silveira Lemos, F.C., Trujillo, E. V., Balanta, P., Macapagal, M.E., Cristina Ferreira, M., Herman, G., de Sauvage, I., Bourguignon, D., Wang, Q., Fulop, M., Harb, C., Chybicka, A., Mekonnen, K.H., Martin, M., Nizharadze, G., Gavreliuc, A., Buitendach, J., Valk, A., Koller, S.H., 2014. Cultural bases for self-evaluation: seeing oneself positively in different cultural contexts. *Pers Soc Psychol Bull* 40, 657–675. doi:10.1177/0146167214522836
- Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.S.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol. Evol.* 24.
- Chapman, R., White, R.G., Shafer, L.A., Pettifor, A., Mugurungi, O., Ross, D., Pascoe, S., Cowan, F.M., Grosskurth, H., Buve, A., Hayes, R.J., 2010. Do behavioural differences help to explain variations in HIV prevalence in adolescents in sub-Saharan Africa? *Trop. Med. Int. Heal.* 15, 554–566.

- d'Albis, H., Augeraud-Véron, E., Djemai, E., Ducrot, A., 2012. The dispersion of age differences between partners and the asymptotic dynamics of the HIV epidemic. *J. Biol. Dyn.* 6, 695–717. doi:10.1080/17513758.2012.688146
- Davison, A., Hinkley, D., 1997. *Bootstrap methods and their application*, Vol. 1. ed. Cambridge University Press.
- de Oliveira, T., Kharsany, A.B.M., Gräf, T., Cawood, C., Khanyile, D., Grobler, A., Puren, A., Madurai, S., Baxter, C., Karim, Q.A., Karim, S.S.A., 2016. Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: a community-wide phylogenetic study. *Lancet HIV* 3018, 1–10. doi:10.1016/S2352-3018(16)30186-2
- Delva, W., Beauclair, R., Welte, A., Vansteelandt, S., Hens, N., Aerts, M., du Toit, E., Beyers, N., Temmerman, M., 2011. Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol. *BMC Public Health* 11, 616. doi:10.1186/1471-2458-11-616
- Delva, W., Meng, F., Beauclair, R., Deprez, N., Temmerman, M., Welte, A., Hens, N., 2013. Coital frequency and condom use in monogamous and concurrent sexual relationships in Cape Town, South Africa. *J Int AIDS Soc* 16, 18034. doi:10.7448/IAS.16.1.18034
- Department, S.D.I. and G., 2012. *City of Cape Town – 2011 Census – Cape Town*. Cape Town.
- Dolezal, C., Marhefka, S.L., Santamaria, E.K., Leu, C.S., Brackis-Cott, E., Mellins, C.A., 2012. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Arch Sex Behav* 41, 401–410. doi:10.1007/s10508-011-9769-6
- Efron, B., 1987. Better bootstrap confidence intervals. *J. Am. Stat. Assoc.* 82, 171–200.
- Gelman, A., 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press, New York.
- Gouws, E., Stanecki, K.A., Lyerla, R., Ghys, P.D., 2008. The epidemiology of HIV infection among young people aged 15-24 years in southern Africa. *Aids* 22, S5–S16. doi: 10.1097/01.aids.
- Gregson, S., Nyamukapa, C.A., Garnett, G.P., Mason, P.R., Zhuwau, T., Carael, M., Chandiwana, S.K., Anderson, R.M., 2002. Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *Lancet* 359, 1896–1903. doi: S0140-6736(02)08780-9
- Hallett, T.B., Gregson, S., Lewis, J.J., Lopman, B.A., Garnett, G.P., 2007. Behaviour change in generalised HIV epidemics: impact of reducing cross-generational sex and delaying age at sexual debut. *Sex Transm Infect* 83 Suppl 1, i50-54. doi: sti.2006.023606
- Harling, G., Newell, M.L., Tanser, F., Barnighausen, T., 2015. Partner Age-Disparity and HIV Incidence Risk for Older Women in Rural South Africa. *AIDS Behav.* doi:10.1007/s10461-014-0952-3

- Harling, G., Newell, M.L., Tanser, F., Kawachi, I., Subramanian, S. V., Barnighausen, T., 2014. Do Age-Disparate Relationships Drive HIV Incidence in Young Women? Evidence from a Population Cohort in Rural KwaZulu-Natal, South Africa. *J Acquir Immune Defic Syndr* 66, 443–451. doi:10.1097/QAI.0000000000000198
- Hastie, T.J., Tibshirani, R., 1990. *Generalised additive models*. Chapman and Hall/CRC, Boca Raton.
- Hawkins, K., Price, N., Mussa, F., 2012. Milking the cow: Young women’s construction of identity and risk in age-disparate transactional sexual relationships in Maputo, Mozambique. *Glob. Public Heal. An Int. J. Res. Policy Pract.* 4, 169–182.
- Hewett, P.C., Mensch, B.S., Erulkar, A.S., 2004. Consistency in the reporting of sexual behaviour by adolescent girls in Kenya: a comparison of interviewing methods. *Sex. Transm. Infect.* 80, ii43-ii48. doi:10.1136/sti.2004.013250
- Kissinger, P., Rice, J., Farley, T., Trim, S., Jewitt, K., Margavio, V., Martin, D.H., 1999. Application of computer-assisted interviews to sexual behavior research. *Am J Epidemiol* 149, 950–954.
- Kuate-Defo, B., 2008. Young People’s Relationships with Sugar Daddies and Young People’s Relationships Sugar Mummies: What do We Know and What do We Need to Know? *Afr. J. Reprod. Health* 8, 13–37.
- Leclerc-Madlala, S., 2008. Age-disparate and intergenerational sex in southern Africa: the dynamics of hypervulnerability. *AIDS* 22 Suppl 4, S17-25. doi: 10.1097/01.aids.
- Luke, N., 2003. Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa. *Stud Fam Plann* 34, 67–86.
- Maughan-brown, B., Evans, M., George, G., 2016. Sexual Behaviour of Men and Women within Age-Disparate Partnerships in South Africa: Implications for Young Women’s HIV Risk. *PLoS One* 1–16. doi: 10.1371/journal.pone.0159162
- Maughan-Brown, B., Kenyon, C., Lurie, M.N., 2014. Partner Age Differences and Concurrency in South Africa: Implications for HIV-Infection Risk Among Young Women. *AIDS Behav.* doi:10.1007/s10461-014-0828-6
- Moore, A.M., Biddlecom, A.E., Zulu, E.M., 2007. Prevalence and meanings of exchange of money or gifts for sex in unmarried adolescent sexual relationships in sub-Saharan Africa. *Afr. J. Reprod. Health* 11, 44–61.
- Morrison-Beedy, D., Xia, Y., Passmore, D., 2013. Sexual risk factors for partner age discordance in adolescent girls and their male partners. *J Clin Nurs* 22, 3289–3299. doi:10.1111/jocn.12408
- Muzondo, I.F., Barry, M., Dewar, D., Whittal, J., Africa, S., Muzondo, I.F., Barry, M., 2004. *Land Conflicts in Informal Settlements: Wallacedene in Cape Town, South Africa*, FIG Working Week 2004. Athens, Greece.
- O’Leary, A., Jemmott III, J.B., Jemmott, L.S., Teitelman, A., Heeren, G.A., Ngwane, Z., Icard, L.D., Lewis, D.A., 2015. Associations between Psychosocial Factors and

Incidence of Sexually Transmitted Disease Among South African Adolescents.  
*Sex Transm Dis* 42, 135–139.

- Ott, M.Q., Barnighausen, T., Tanser, F., Lurie, M., Newell, M.L., 2011. Age-gaps in sexual partnerships: seeing beyond 'sugar daddies'. *AIDS* 25, 861–869.
- Penone, C., Davidson, A.D., Shoemaker, K.T., DiMarco, M., Rondinini, C., Brooks, T.M., Young, B.E., Graham, C.H., Costa, G.C., 2014. Imputation of missing data in life-history trait datasets: which approach performs best? *Methods Ecol. Evol.* 5, 961–970.
- Pettifor, A.E., Rees, H. V, Kleinschmidt, I., Steffenson, A.E., MacPhail, C., Hlongwa-Madikizela, L., Vermaak, K., Padian, N.S., 2005. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. *AIDS* 19, 1525–1534. doi:00002030-200509230-00012 [pii]
- Phaswana-Mafuya, N., Shisana, O., Davids, A., Tabane, C., Mbelle, M., Matseke, G., Banyini, M., Kekana, Q., 2014. Perceptions of sugar mommy practices in South Africa. *J. Psychol. Africa (south Sahara, Caribbean, Afro-Latin Am.* 24, 257–263. doi:10.1080/14330237.2014.906086
- R Development Core Team, 2014. R: A language and environment for statistical computing.
- Ritchwood, T.D., Hughes, J.P., Jennings, L., MacPhail, C., Williamson, B., Selin, A., Kahn, K., Gómez-Olivé, F.X., Pettifor, A., 2016. Characteristics of Age-Discordant Partnerships Associated with HIV Risk Among Young South African Women (HPTN 068). *JAIDS J. Acquir. Immune Defic. Syndr.* 72, 423–429. doi:10.1097/QAI.0000000000000988
- Rubin, D., 1987. *Multiple Imputation for Nonresponse in Surveys*. Wiley, New York.
- Rubin, D.B., 1976. Inference and missing data. *Biometrika* 63, 581–592. doi: <https://doi.org/10.1093/biomet/63.3.581>
- Schaefer, R., Gregson, S., Eaton, J.W., Mugurungi, O., Rhead, R., Takaruzza, A., Maswera, R., Nyamukapa, C., 2017. Age-disparate relationships and HIV incidence in adolescent girls and young women: evidence from a general-population cohort in Zimbabwe. *AIDS*. doi:10.1097/QAD.0000000000001506
- Shah, A.D., Bartlett, J.D., Carpenter, J., Nicholas, O., Hemingway, H., 2014. Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study. *Am J Epidemiol* 179, 764–774.
- Shisana, O., Rehle, T., Simbayi, L., Zuma, K., Jooste, S., Zungu, N., Labadarios, D., Onoya D, 2014. South African National HIV Prevalence, Incidence and Behaviour Survey, 2012. Cape Town.
- Smith, J., Nyamukapa, C., Gregson, S., Lewis, J., Magutshwa, S., Schumacher, C., Mushati, P., Hallett, T., Garnett, G., 2014. The Distribution of Sex Acts and Condom Use within Partnerships in a Rural Sub-Saharan African Population. *PLoS One* 9, e88378. doi: 10.1371/journal.pone.0088378

- Stekhoven, D.J., Buhlmann, P., 2012. MissForest - nonparametric missing value imputation for mixed-type data. *Bioinformatics* 28, 112–118.
- Street, R.A., Reddy, T., Ramjee, G., 2015. The generational effect on age disparate partnerships and the risk for human immunodeficiency virus and sexually transmitted infections acquisition. *Int J STD AIDS*. doi:10.1177/0956462415592325
- van Buuren, S., Groothuis-Oudshoorn, K., Buuren, S., Groothuis-Oudshoorn, K., 2011. MICE: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 45, 1–67.
- Weiser, S.D., Leiter, K., Bangsberg, D.R., Butler, L.M., Percy-de Korte, F., Hlanze, Z., Phaladze, N., Iacopino, V., Heisler, M., 2007. Food insufficiency is associated with high-risk sexual behavior among women in Botswana and Swaziland. *PLoS Med* 4, 1589–97; discussion 1598. doi: 10.1371/journal.pmed.0040260
- White, I.R., Royston, P., Wood, A.M., 2011. Multiple imputation using chained equations: Issues and guidance for practice. *Stat. Med.* 30, 377–399. doi:10.1002/sim.4067
- Wood, S., 2006. *Generalized Additive Models: An Introduction with R*. Taylor & Francis.

## Supplementary Web Appendix 1: Checking the Missing at Random (MAR) assumption

A particular variable is likely to be considered Missing at Random (MAR) when observations have a high probability of being missing because they have particular values of some other characteristics. Thus, we tested to see if our key variables (i.e. Partner ages and HIV infection status) had different distributions of participant characteristics, based upon their missingness status (See Tables 1 and 2 below). Data are considered to be Missing Not at Random (MNAR) when the likelihood that the observation is missing depends on the true value of that variable. Unfortunately, it is not possible to tell the difference between MAR and MNAR missingness mechanisms by examining the data. Therefore, we made our best assessment of whether we believed the data was MAR or MNAR based upon our *a priori* knowledge of the likely mechanism.

For partner ages, we believe that the MAR assumption is accurate because we used Audio Computer-Assisted Self-Interviewing (ACASI) to collect the survey data. This method has been demonstrated to reduce social desirability bias (SDB) in reporting sensitive sexual behaviours. Thus, we do not think participants would have been likely to skip the question about their partner's age, based upon the age itself.

There is a chance that the HIV status data may be MNAR, since participants who previously tested positive may have been less likely to consent to HIV testing in the ZAMSTAR study.

**Table 1. Comparison of participant characteristics for relationships with missing partner ages versus non-missing ages**

Variables	Partner age	
	Missing No. 65	Observed No. 805
Age		
Median (IQR)	34.0 (29.0 - 49.0)	29.0 (24.0 - 42.0)
Missing	4 (6.2%)	19 (2.4%)
Sex		
Male	46 (70.8%)	323 (40.1%)
Female	19 (29.2%)	455 (56.5%)
Missing	0 (0.0%)	27 (3.4%)
HIV status		
Negative	36 (55.4%)	442 (54.9%)
Positive	18 (27.7%)	201 (25.0%)
Missing	11 (16.9%)	162 (20.1%)
Employed		
Yes	7 (10.8%)	205 (25.5%)
No	57 (87.7%)	598 (74.3%)
Missing	1 (1.5%)	2 (0.2%)
Highest grade completed		
Primary	35 (53.8%)	221 (27.5%)
Secondary	29 (44.6%)	529 (65.7%)
Tertiary	1 (1.5%)	53 (6.6%)
Missing	0 (0.0%)	2 (0.2%)
ACASI survey difficulty		
Very easy	43 (66.2%)	558 (69.3%)
Somewhat easy	10 (15.4%)	137 (17.0%)
Somewhat difficult	10 (15.4%)	95 (11.8%)
Very difficult	2 (3.1%)	15 (1.9%)
Relationships in past year	3.0 (2.0 - 11.0)	2.0 (1.0 - 4.0)
Partner type		
Casual	36 (55.4%)	166 (20.6%)
Main	29 (44.6%)	639 (79.4%)

All continuous values are reported with median and inter-quartile range, Med (IQR), while categories are reported in percentages, n (%).

**Table 2. Comparison of participant characteristics for those with missing HIV status versus non-missing statuses**

Variables	HIV status	
	Missing	Observed
	No. 90	No. 416
Age		
Median (IQR)	31.0 (25.0 - 39.0)	34.0 (26.0 - 44.0)
Missing	3 (3.3%)	14 (3.4%)
Sex		
Male	33 (36.7%)	130 (31.2%)
Female	56 (62.2%)	274 (65.9%)
Missing	1 (1.1%)	12 (2.9%)
Employed		
Yes	26 (28.9%)	92 (22.1%)
No	64 (71.1%)	322 (77.4%)
Missing	0 (0.0%)	2 (0.5%)
Highest grade completed		
Primary	23 (25.6%)	133 (32.0%)
Secondary	62 (68.9%)	263 (63.2%)
Tertiary	5 (5.6%)	18 (4.3%)
Missing	0 (0.0%)	2 (0.5%)
ACASI survey difficulty		
Very easy	60 (66.7%)	289 (69.5%)
Somewhat easy	13 (14.4%)	65 (15.6%)
Somewhat difficult	15 (16.7%)	52 (12.5%)
Very difficult	2 (2.2%)	10 (2.4%)
Relationships in past year	1.0 (1.0 - 2.0)	1.0 (1.0 - 2.0)

**All continuous values are reported with median and inter-quartile range, Med (IQR), while categories are reported in percentages, n (%).**

## Supplementary Web Appendix 2: Variables included in the imputation model

We formulated the imputation model by including a combination of different socio-demographic variables and relationship characteristics that may have been predictive of the missing values. Additionally, the imputation models included all of the variables used in subsequent analytical models, as well as meta-questions, which were related to use of the ACASI computer survey. We hypothesized that if the participants found the ACASI computer survey to be difficult, for example, they may be more likely to skip questions.

The imputation models were specified according to the procedures outlined in (van Buuren et al., 2011). Thus, these variables were used:

- Age
- Sex
- Employment status
- Highest level of education
- Number of relationships
- Partner type
- Partner's place of residence
- Whether participant slept with that person in the previous year
- Whether participant thought the partner had other partners
- Whether relationship was ongoing at the time of the survey
- Concurrency status of relationship
- Sex frequency during relationship
- Condom use
- Relationship duration
- HIV status of participant
- Age of partner
- Whether participant thought the interviewing modality was easy
- Whether participant thought they had privacy while doing the interview
- Whether participant claimed to answer truthfully
- Participant's preferred interviewing modality

## References

van Buuren, S., Groothuis-Oudshoorn, K., Buuren, S., Groothuis-Oudshoorn, K., 2011. MICE: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 45, 1–67.

### Supplementary Web Appendix 3: Relationship characteristics and complete case analysis

**Table 1. Relationship characteristics by sex of participant using the original dataset.**

Variables	Sex	
	Male No. 369	Female No. 474
Age		
15-24	106 (28.7%)	125 (26.4%)
25-34	106 (28.7%)	167 (35.2%)
35-44	57 (15.4%)	99 (20.9%)
45-70	93 (25.2%)	67 (14.1%)
Missing	7 (1.9%)	16 (3.4%)
Employed		
Yes	120 (32.5%)	85 (17.9%)
No	247 (66.9%)	388 (81.9%)
Missing	2 (0.5%)	1 (0.2%)
Highest grade completed		
Primary	119 (32.2%)	126 (26.6%)
Secondary	211 (57.2%)	332 (70.0%)
Tertiary	39 (10.6%)	14 (3.0%)
Missing	0 (0.0%)	2 (0.4%)
HIV status		
Negative	240 (65.0%)	219 (46.2%)
Positive	46 (12.5%)	166 (35.0%)
Missing	83 (22.5%)	89 (18.8%)
Number relationships in past year	3.0 (2.0 - 5.0)	1.0 (1.0 - 3.0)
Partner age		
<= 24	143 (38.8%)	55 (11.6%)
25-34	89 (24.1%)	170 (35.9%)
35-44	51 (13.8%)	113 (23.8%)
>= 45	40 (10.8%)	117 (24.7%)
Missing	46 (12.5%)	19 (4.0%)
Partner type		
Casual	136 (36.9%)	53 (11.2%)
Main	233 (63.1%)	421 (88.8%)
Relationship was concurrent		
No	93 (25.2%)	269 (56.8%)
Yes	209 (56.6%)	163 (34.4%)
Missing	67 (18.2%)	42 (8.9%)
Sex frequency		
None	7 (1.9%)	8 (1.7%)
1 time/week	69 (18.7%)	142 (30.0%)
2 times/week	71 (19.2%)	149 (31.4%)
3 times/week	59 (16.0%)	75 (15.8%)
4 times/week	27 (7.3%)	21 (4.4%)
>= 5 times/week	32 (8.7%)	35 (7.4%)
Missing	104 (28.2%)	44 (9.3%)
Condom frequency		
Never/inconsistent	194 (52.6%)	263 (55.5%)
Always	118 (32.0%)	173 (36.5%)
Missing	57 (15.4%)	38 (8.0%)

All continuous values are reported with median and interquartile range, Med (IQR), while categories are reported in percentages, n (%). There were 27 relationships that belonged to participants with missing gender status, and thus were excluded from this table.

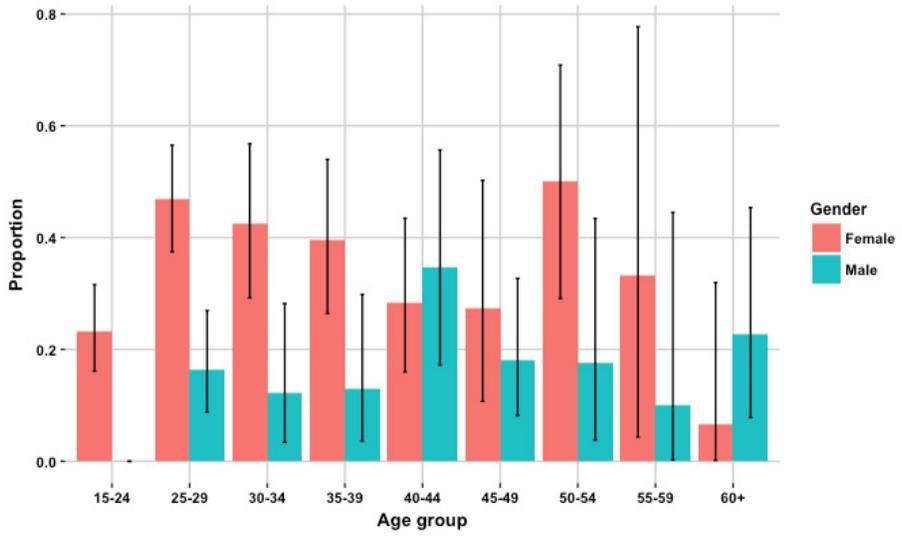


Figure 1. HIV prevalence among complete cases, with 95% exact confidence intervals

**Table 2. Linear mixed effects model age-mixing pattern features for different population strata.** This analysis was done with complete cases.

Population strata	B-coefficient	Intercept	BSSD	WSSD
HNW	0.81 (0.68 — 0.94)	23.65 (20.52 — 26.78)	9.94 (8.67 — 11.25)	3.71 (3.03 — 4.69)
HPW	0.70 (0.54 — 0.86)	24.28 (21.06 — 27.5)	0.00 (0.00 — 3.39)	9.95 (8.89 — 11.08)
HNM	0.63 (0.44 — 0.82)	19.57 (14.82 — 24.31)	12.50 (10.73 — 14.4)	3.15 (2.79 — 3.60)
HPM	0.66 (0.20 — 1.13)	21.30 (6.83 — 35.77)	9.74 (3.36 — 13.77)	4.93 (3.31 — 9.29)

HNW, HIV Negative Women  
 HNM, HIV Negative Men  
 HPW, HIV Positive Women  
 HPM, HIV Positive Men  
 BSSD, Between-subject Standard Deviation  
 WSSD, Within-subject Standard Deviation

**Table 3. Model coefficients for relationship between HIV and bridge width.** The Expected Bridge Width Ratio (EBWR) is the ratio of expected bridge widths for those who are HIV positive compared to those who are HIV negative. The EBWR is derived from the negative binomial model that also adjusts for the age of the participant.

Analysis type	Men EBWR (95% CI)	Women EBWR (95% CI)
Complete case analysis	1.95 (0.62 — 6.13)	3.61 (1.42 — 9.15)
Pooled from MICE	1.93 (0.69 — 5.41)	2.76 (1.16 — 6.59)

**The complete case analysis uses the raw dataset and observations for which there were no missing data on variables used in the analysis (Women: n = 56; Men: n = 58). The pooled MICE analysis uses the combined 100 imputed datasets presented in the main text.**

## CHAPTER 8: DISCUSSION

---



Photo credit: Jonathan Morgan via Visualhunt / CC BY-NC-ND

## 8.1 CHAPTER PRECIS

---

At the commencement of this dissertation, I aimed to gain an increased understanding about the role that age-mixing might play in maintaining the HIV epidemic in Southern Africa and whether age asymmetries have the potential to increase individual-level risk of HIV acquisition. In the section that follows, I will summarize the key findings of the research reported in Chapters 3 to 7, and elaborate on HIV epidemiological insights they provide. Following this, I will discuss the over-arching limitations of my thesis research, and then end the chapter by providing some research recommendations that can be derived from my key results.

## 8.2 EPIDEMIOLOGICAL IMPLICATIONS OF KEY FINDINGS

---

### 8.2.1 Risky sexual behaviours correlated with large age differences

In the CAPS, LNS, and SIHR analyses, we elucidated sexual behaviours that were associated with age asymmetries. In each of those studies, we found that large age asymmetries were correlated with reductions in condom use. In the SIHR study, we reported that larger age differences resulted in higher sex frequencies. In the same vein, the CAPS study we found that cohabiting with a partner, which has been linked to higher sex frequencies (175,176), was also associated with age differences. Additionally, we demonstrated that young women in the SIHR study who had the largest age differences also had the longest relationship durations. Young women in the CAPS population who were in an ADR reported more STI symptoms than those who were not in ADR. These findings are in agreement with the causal pathways previously outlined in Section 1.3.

The probability that an HIV negative individual will become infected is largely a function of the amount of exposure that individual has to HIV and the duration of those exposures (177). In my analyses, the increased reporting of STI symptoms, reduced condom use, higher sex frequencies, and longer relationship durations from women in relationships with large age differences implies that they would have a higher probability of acquiring HIV because their relationships would have provided more exposure opportunities to HIV.

In the LNS we also found that men who reported that they had a concurrent relationship in the past three years also had larger bridge widths during that time. Serial monogamy is thought to “trap” HIV within a couple until the relationship dissolves. Concurrence, on the other hand, allows the virus to be transmitted more quickly because, by definition, the relationship does not need to end before a new person can be infected (178,179). Since a person’s viral load also peaks in the first few months of infection, that person would be more likely to pass HIV on to a seronegative individual outside of the relationship from which they acquired the infection (46,180). In the LNS population concurrency may not only provide the potential for an epidemic within a specific age cohort, but since it is also associated with large bridge widths, concurrency may also facilitate endemic HIV by speeding up the rate of infection between age cohorts.

### 8.2.2 Age-mixing may create conditions necessary to sustain HIV in the population

One of the main objectives of my research was to summarize key features of age-mixing patterns in different populations of South Africa and Malawi. In the representative sample of young adults from the Cape Town CAPS population, we observed that as the ages increased for young adult men their age differences also tended to increase. Whereas, for women the age differences started off larger, compared to the young men, and they remained relatively constant for older ages. In the LNS population of Malawi, we witnessed the same pattern. However, in the CTSBS population men were initially younger than their partners, on average, and then as their ages increased, they gradually became older than their female partners. Among the women, at younger ages they had the largest age differences, and the age gaps narrowed as they grew older.

Though both the CTSBS and CAPS took place in Cape Town there are at least a few reasons they might have different age-mixing patterns. The first is that the participant age windows observed in the studies were different. In the CAPS analysis, we only observed participants that were 17-26 years old, versus the CTSBS population, which had participants aged 15-70 years old. Moreover, the CTSBS study population was composed of individuals from high HIV prevalence settings, whereas CAPS included a representative sample from the entire Cape Town metropolitan. This difference may also provide insight as to why the HIV prevalence is so much higher in the CTSBS population than the general Cape Town population. Both young men and women were choosing older partners of the opposite sex, who had higher background prevalences of HIV infection.

Interestingly, in the LNS population we found that women had larger between- and within-individual variability in partner ages than men. While this was not precisely quantified in the CAPS population, we also observed that black women, in particular, had larger variability at younger ages, and this variability remained large at all ages. This pattern of increasing variability in partner ages (or age differences) for men, and a constant but relatively high variability for women, was also observed in a cohort study from Zimbabwe (181). In the CTSBS population the HIV positive women had the largest within-individual variability out of all the groups, and we could detect meaningful difference between them and the HIV negative women. In general, few other studies in Southern Africa have attempted to visualize or quantify the age-mixing pattern the way we have in our analyses (181,182).

The reason for women having large variability of partner ages is context specific. On Likoma Island, for example, at the time of the initial study, some young women had been noted to spend time by a military camp on the island where there were young, well-paid soldiers staying (129). These soldiers had often been blamed for unwanted pregnancies and STIs on the island. For some women, these soldiers may have represented short or once-off sexual relationships with someone closer to their own age, before they married an older, more settled man. Moreover, on Likoma island, there is a local sexual practice called “kutondola”, which may have introduced some within-subject variation in partner ages for men (129). Historically, the practice allowed older, unmarried women to “rent” a neighbour or friend to help her get pregnant. These arrangements were often done in secret and the identity of the man was often unknown to the woman herself. This practice may have introduced dynamics into the population whereby some female partners were older than their male “kutondola” partners, while the male partners were still married to their younger female spouses.

In some urban settings of Southern Africa, women have been noted to engage in “third party patronage” relationships (14), which also may explain the large within-subject variation we witnessed among women in Cape Town. The young women typically begin relationships with older men in order to acquire resources, such as money for living expenses, tuition, clothing, etc. The acquisition of these resources help the younger women to become more attractive partners for age-similar men whom they wish to marry (25,93,183).

Large variances in partner ages may be the key to the endemicity of HIV in various Southern African populations. For instance, if the variance in partner ages had been low among the first age cohorts of women that were infected with HIV, then they would have only been able to transmit the virus to the men within their age cohort, creating a temporary epidemic that would have been eliminated as those individuals aged and died (98). However, the large variability we witnessed in older men and among women of all ages in the populations we studied, means that HIV transmission could continually occur between different age cohorts, thus allowing HIV to become endemic in those regions.

### **8.2.3 Relationship between age differences and prevalent HIV infection in individuals**

Two of our analyses facilitated the study of the association between age difference indicators and prevalent HIV status. We observed that the 18 to 49-year-old women in the LNS had increasing probabilities of being HIV positive as their maximum age differences increased from 2 to 12. When the largest age differences were greater than 12, the probability of being HIV positive decreased. This may indicate that women could be at increased risk of acquiring HIV from some of their older partners, though if the partner is more than 12 years older, then women on Likoma Island may actually be protected from HIV transmission. This finding reinforces the idea, that arbitrary categories for age differences may mask true underlying associations between HIV infection and age asymmetries. In some contexts, such as Likoma Island, a prescription for young women to avoid all relationships with men who are five or more years older, may miss opportunities to prevent HIV transmission by encouraging partnering with older, safer men.

Men in the LNS that were 18 to 49-years-old, who were on average six or more years younger than their partners, had nearly five times the odds of being HIV positive, compared to those who were 0-4 years older than their partners. This indicates that the discourse around age-mixing needs to expand to incorporate young men, as they are also clearly at risk of acquiring HIV from women who are older than them in some populations. Moreover, the CTSBS study also showed that young men tend to choose older female partners. If, for example, a man in the 15-24-year age group chose a partner from the CTSBS population who was in the 25-29-year group, where the HIV prevalence was approximately 45%, they would have had a greater chance of selecting an HIV positive partner compared to if they had selected a woman from their own age group where the HIV prevalence was 31%.

In the CTSBS study, we also saw that women who were HIV positive had larger expected bridge widths, than the HIV negative women. It is important to remind readers that these bridge widths were calculated from relationships that occurred after the participant's HIV diagnosis. While we do not know which partner – if any from the

ones enumerated – may have caused the women to acquire HIV, the fact that those with HIV are still choosing a wide range of partners implies that there may be opportunities for the women to transmit HIV to different age cohorts during unprotected intercourse.

These findings emphasise the point that HIV transmission is a two-way street: young women maybe be at risk of acquiring HIV from older men, but the older men who are already HIV positive likely contracted HIV from age-similar women when they were younger. Interventions that attempt to extinguish transmission pathways, should consider both sides of the equation (See Section 9.1).

#### **8.2.4 Psycho-social determinants of age differences in relationships**

In the analysis of 17 to 26-year-old women from the CAPS study, we identified a few psycho-social predictors of ADR and large age differences in relationships. We found that among black women, spending more nights in the same household per week and having both parents still alive were protective of being in an ADR. This points to the importance of having a relatively stable home and family life. Indeed, in the CTSBS qualitative study one woman in particular, stated that women involved in ADR might be looking for father-figures. In general, several women articulated that these young women might desire someone who is willing to look after them, give them attention, and provide pocket-money. Those qualitative insights explain one of the other key results from the CAPS data: women in ADR were more likely to have accepted monetary or in-kind support from that partner. Other studies from urban contexts in Southern Africa have found similar results (32,97,184).

From the perspective of women in the CTSBS qualitative study (whose ages ranged from 18 to 65-years-old), there were few risks to engaging in in ADR. On the contrary, age-similar partnerships were deemed to be riskier because they were more likely to involve abuse and disrespect. Our qualitative study was not able to confirm the truth of these perceptions, and there is conflicting evidence about whether intimate partner violence is associated with large age differences in relationships (185–190). Like most inferences, those associations are highly contextual. Regardless of whether age-similar partnerships are more likely to involve violence, negative perceptions of these relationships in certain contexts may encourage women to form ADR, which may involve riskier sexual behaviours (See Section 8.2.1). The psycho-social factors that motivate women to begin these relationships are important to identify, because they have implications for whom to target when developing HIV sexual behaviour risk-reduction interventions.

#### **8.2.5 The role of structural interventions in reducing age differences**

At the end of the manuscript in Chapter 3, we advocated for HIV prevention interventions that take into account psycho-social motivations young women have for entering into ADR. Poverty, unstable homes, absent parents, and food insecurity are all likely to influence a woman's decision to begin a relationship with an older man despite the HIV-related risks these relationships may pose. Since that manuscript's publication in 2012 several structural and behavioural interventions have been conducted, which aim to retain young women and adolescents in school, as well as improve long-term economic opportunities (131,134,191–193). It has been hypothesized that a side-effect of these interventions might be that the young women engage in less risky sexual behaviours, such as ADR, and thereby remain HIV negative (194–196). Indeed, we saw

some evidence that this might be at work in the CAPS population, for which the black women who attended more classes, were also less likely to be in an ADR.

The SIHR study, which was the focus of Chapter 6, was one such structural intervention. Its initial results that were published while the cash transfer programme was still taking place found that girls in the intervention groups had lower HIV prevalence at the 18-month follow up, and the girls were also less likely to report a relationship with a man 25 years or older (131). We hypothesized that the intervention would have reduced age differences in relationships, and thereby exposure to HIV through it. While we did, indeed, find that large age differences were associated with lower levels of condom use, increased sex frequency, and longer relationship durations (See Section 8.2.1), the intervention itself had only a small effect on age differences. Our findings were not statistically significant, but they were in the hypothesized direction: intervention groups had smaller age differences. This is similar to a study of government-administered cash transfer programmes in South Africa, which found that child support grant reduced the prevalence of ADR (197).

One intriguing finding from the SIHR analyses was that the risk behaviours seemed to be reduced in all study groups while the intervention was ongoing. We attributed this finding to the Hawthorne Effect (198,199). In other words, we believe that the act of taking interest in all girls and their participation in the study, regardless of the study group, may have prompted them to engage in (or report) less risky sexual behaviours. Once the intervention was done, and cash was no longer forthcoming, the girls of all study groups may have perceived a loss of interest in them and thus lost the motivation to engage in safer sexual behaviours.

The principle investigators of the SIHR also found that two years after the cash transfer programme there were few sustained positive outcomes in the medium-term, and they even went so far as to claim that the interest in cash transfers “may be misguided” (113, p.1). In order for cash transfer programmes to have long-term effects, it has been argued that they need to translate into the accumulation of productive assets and vocational skills, which will increase future earnings (200,201). It is possible that the amount of money that was transferred to the girls was enough to keep them from older men while the intervention was ongoing, but the cash was not enough to create ongoing revenue streams after the programme had ended.

### 8.3 LIMITATIONS

---

The primary limitation that cross-cut each study presented here was our restricted ability to measure sexual behaviours accurately. In our studies, as well as most public health studies, sexual behaviour was self-reported in either surveys or interviews. Self-reports of sexual behaviour are typically prone to under-reporting if the behaviour happens to be stigmatized in that particular setting, or over-reported if the behaviour is normative (202). Moreover, recall bias often occurs, meaning the participants are more likely to recall certain behaviours (e.g. inconsistent condom use) if they already have the outcome of interest (e.g. HIV positive) (203,204). Unlike other health risk behaviours, sexual behaviours are not directly observable. Prostate-specific antigen is a physiological biomarker that can give an indication of whether a woman engaged in unprotected sex during the previous two days (205). However, there are no biomarkers that can demonstrate the frequency of sex or associated behaviours (202). Due to

practical and ethical issues related to direct measurement and observation of sexual behaviours, self-reported data continues to be the norm (202).

In order to mitigate some of these biases, some of the studies took preventative measures in the study design and data collection phases of the projects. The LNS and CTSBS studies used Audio Computer-Assisted Self-Interviews (ACASI) for the questionnaire, as opposed to face-to-face-interviews (FTFI), which are commonly used in public health and demographic studies. It has been theorized that ACASI lends itself to reporting more sensitive and stigmatized behaviours because participants in those studies are guaranteed privacy and anonymity, unlike in FTFI, which can lead to social desirability bias. Indeed, several studies have found this to be the case in different African contexts (118,120,121,206). We also conducted a study with the CTSBS data to find out if ACASI would be suitable for our study setting and if it resulted in improved accuracy (113, Appendix A). Not only did most of the participants find the ACASI questionnaires to be easy to use, confidential, and preferable to other modes of interviewing, but the study also found there to be less social desirability bias present than FTFI used in DHS questionnaires in Southern African settings. While these findings cannot guarantee that all of the sexual behaviours were accurately measured, we feel confident that the LNS and CTSBS studies had less bias.

With the CTSBS survey we also took additional initiatives to improve accuracy of reporting. The first is that the questionnaire displayed a visual timeline on the screen to help participants keep track of when different relationships occurred in the previous year. Visual cues for anchor dates, such as timeline markers for start and end dates of relationships, have been demonstrated to result in more internally consistent reports of relationship histories (207,208). The second action we took was to keep the overall survey relatively short. We only asked a few pertinent questions at the episode-level, relationship-level, and participant-level. This, we believe, helped to reduce fatigue bias. Fatigue bias is a form of non-response bias that may result in participants leaving questionnaire items missing, or refusing to participate simply because they do not want to take the time to answer the questions (209). In sexual behaviour surveys this can happen if, for example, many questions are asked for each sexual partner, leading to a participant only nominating one partner, when they may have had many more. Another way this occurs is when a participant is exposed to many long interviews (for example, DHS surveys are notorious for containing many long modules), over several visits. After the first visit the participant may be “test wise” and refuse to participate in subsequent visits. We think fatigue bias may have been present in both the CAPS study and the SIHR study, which had several follow-ups and many long modules.

In the CTSBS qualitative study we tried to elicit truthful and thoughtful responses by conducting the interviews in private environments away from the listening ears of friends or family. Moreover, we tried to put participants at ease by conducting the interview in the language of their choice (i.e. isiXhosa, Afrikaans, or English). This, coupled with the fact that both my research assistant and I were female and not resident in the study communities, hopefully established the feeling that their views and responses would not be shared with others in their community.

A final limitation to all of the quantitative analyses presented in this dissertation was that participant and partner ages may have also been inaccurately reported. In several areas of the world people do not necessarily mark the passage of time in years since their birth (210). Thus, participants in demographic and public health studies from

these settings may guess at their age and this can result in age heaping (210–215). Age heaping occurs when there is a systematic preference or rounding of ages, oftentimes to the nearest age that ends in ‘5’ or ‘0’.

There is some evidence that participants in the LNS study may have misclassified their partners’ ages in the survey (216). However, we mitigated this limitation by cross-referencing the given ages with the self-reported ages of the partners, where they were available. In the CTSBS survey, we tried to diminish the effect of partner age misclassifications by first prompting participants to identify the year their partner was born. If they did not know the year, they could alternatively report their partner’s age. In the U.S. context, this has been shown to be more accurate than directly asking about age (210). Unfortunately, it is not known if the year of birth can provide a more accurate age in different African contexts, or with partner ages, specifically. With both of the South African studies in this dissertation, we have reason to believe there is probably fairly accurate reporting of the participant’s own age due to the dedicated vital statistics programme in South Africa, and the country’s policy of issuing identification books to all citizens (217). It is a shortcoming of my study that we do not have much information on the relative accuracy of ages reported by girls in the SIHR study.

## 8.4 RESEARCH RECOMMENDATIONS

---

### 8.4.1 Improvement in the design of sexual behaviour studies

Given the limitations of my research that I enumerated in the preceding section, one of the primary ways in which the public health community can improve the study of age-mixing dynamics is in advocating for more well-designed, longitudinal sexual behaviour studies that also measure HIV biomarkers. Currently, we exist in an economic and research climate that encourages large-scale studies of biomedical interventions to curtail the spread of HIV, while behaviour research is relatively under-funded (218). Large-scale, multi-million dollar studies are very good at assessing group and time effects of interventions on HIV incidence, while their studies of relationship histories are of secondary importance (219), and thus they use sub-optimal designs for measuring this information. Conversely, studies with a primary focus on measuring sexual behaviours well, rarely are funded to collect HIV biomarkers (220).

In more ideal circumstances sexual behaviour studies should collect data on sexual behaviours and HIV status over time, in order to determine the temporality of the associations of sexual behaviours with HIV infection, as well to assess whether behavioural dynamics are changing over time. Above and beyond this, the studies need to make great efforts to reduce the biases discussed above, either through improvements to the interviewing modality, such as using ACASI, and through improving the measures themselves. Some ways in which to improve the accuracy of measured sexual behaviours are to: 1. use psychometrically evaluated measures, 2. validate and standardize the measures, 3. use language that is easy to understand, 4. use measures appropriate for purpose, 5. pilot the questionnaires in order to probe for sensitivities and cultural issues, 6. implement techniques that improve recall, such as, timelines and anchor dates, 7. establish rapport and trust with participants and communities, 8. ask questions in a direct way without apology, 9. sequence questions from least to most threatening, and 10. be sensitive to contextual issues (202).

It is also vitally important to capture information about as many relationships as possible in the previous year(s). Importantly, if variation in partner ages is an important determinant of sustained HIV transmission in a population, as we argue, then researchers need to know about all relationships occurring in a given amount of time, regardless of the partnership type. This is especially critical in contexts of high HIV prevalence, where partner concurrency is also common (40,43,221,222). In other words, it is not sufficient to study age-mixing dynamics with only the “most recent” or “primary” partnership, as so many sexual behaviour studies have done. In order to capture all the relationships that occurred within the previous year, it is central keep the number of questions short and relevant. Asking too many questions may result in truncated reports of partners.

#### **8.4.2 Analytical recommendations**

I can recommend four different strategies to improve our understanding of age-mixing and HIV epidemics, which should take place during the analytical phase of studies. First, researchers should shift their focus away from studying age-disparate relationships, dichotomized as relationships where the male partner is 5 or more years older than the female partner. As pointed out in Section 2.3.3, categorizing continuous covariates runs the risk of missing natural breakpoints in the variable where changes in the risk of HIV infection (or other sexual behaviour outcomes) actually occur. I believe this approach has partially been responsible for the absence of evidence for a causal association between age differences and HIV incidence. A better tactic would be to treat age difference as a continuous covariate that is allowed to have a flexible non-linear relationship with the outcome of interest. For this, splines and generalized additive models would be a more suitable approach to studying these dynamics. After the publication of Chapter 5, others have begun to analyse age differences in similar ways (223). To add to this, more research needs to be done to determine the best age difference indicators, which are able to predict the likelihood of an individual having HIV infection. As illustrated in this dissertation, maximum age difference in a given period of time is one such example that may have the potential to be more predictive of HIV infection risk.

Secondly, studying participant-level risk of HIV infection for different indicators using flexible models in sex-stratified groups, requires very large sample sizes of participants. The study of bridge widths, in particular, requires many participants to report more than one relationship. Often, this is not feasible with real-world populations due to the constraints outlined in the previous section. One potential way forward is to conduct stochastic simulation studies, calibrated to real-world data. Then researchers can use this simulated data to perform similar analyses as I did here using the techniques outlined above.

The third analytical recommendation is for more population-level analyses of age-mixing patterns in different contexts. The population-level indicators: rate of partner age change for each year increase in age of participant, the average partner age among the youngest sexually active people in a population, as well as the between and within-subject variability in partner ages should be calculated and compared in different settings. For example, comparisons of the age-mixing pattern features observed in my study populations with epidemics in Western European Immigrant populations, prison populations, or even homosexuals in Southern Africa, may improve our understanding

of why population HIV prevalence and incidence are so high in different epidemic settings as well as, age and sex stratified groups within those settings.

Finally, in the LNS and CTSBS analyses, we assumed the missing partner age and HIV data was Missing at Random (See Section 2.3.4), and thus justified the use of MICE techniques for imputation. However, there is a possibility that this data was Missing Not at Random, and thus MICE may not have eliminated all of the bias that would have resulted (although it probably helped, regardless (164)). One potential avenue for future research would be to conduct various sensitivity analyses in this population to see how the results would change under different assumptions about partner age missingness. For example, one could assume that the larger the age difference in a relationship, the more likely the partner age was to be missing, and thus take that into account in the imputations of partner age. The study of using MICE for MNAR data is still in its infancy (224).

## CHAPTER 9: CONCLUSIONS

---



Photo credit: Alexis Strimenos

## 9.1 RECOMMENDATIONS FOR PUBLIC HEALTH INTERVENTIONS

---

In Southern Africa, and especially recently in South Africa, anti-sugar daddy campaigns trying to prevent ADR among young women, have been commonplace (225). Politicians have been openly discouraging them and blaming them for the HIV epidemic in South Africa (226,227), though evidence that ADR actually cause HIV infection have been limited. In 2014 when the widely publicized Harling et al. study did not find any evidence of ADR driving HIV incidence (108), many organizations and activists were very quick to stop advising against ADR (228). For the past few years, the HIV prevention community has been in abeyance about whether ADR-prevention interventions are appropriate, and if so, what should the messages contain.

The results from the studies presented in this dissertation offer a more nuanced picture about the nature of age differences and HIV transmission, and suggest that moralizing approaches stigmatizing all relationships where the male partner is 5 or more years older, may not be appropriate, nor evidence-based. In reality, there is probably not some magical cut-off for age differences that confer risk across all populations. Moreover – though this is common sense – we must not forget that age differences have the ability to affect transmission risk in men, as well as women. Complicating all of this, is the observation that high individual variability in partner ages, in addition to relationship age differences, is probably what drives transmission. Given these complexities, I propose six public health intervention recommendations:

1. Raising awareness about potential HIV risks associated with larger age differences should be context dependent and evidence-based. For example, based upon the research presented here, it might be appropriate to raise awareness among 18-49-year-old women on Likoma Island about the fact that increasingly large age differences may put them at increased risk of HIV infection, and potentially protect them if their partners are substantially older. It would not be appropriate to use that same information to raise awareness about age difference related risks in an urban, Johannesburg setting. It should be additionally noted that the results from one study alone should not be used as the basis for an HIV risk awareness campaign. Most of the studies presented in this dissertation are exploratory and should be repeated in different populations and for stratified age groups.
2. As indicated in Chapter 4, women may perceive age-similar partners to be potentially more abusive than age-disparate partners. It is not clear whether this perception by our Cape Town interlocutors is accurate or based on any evidence. A recent study utilizing CAPS data, which examined the accuracy of HIV-risk perceptions among young people, found that HIV-risk perceptions are often quite inaccurate in this context (229). It is entirely possible that if their perceptions are incorrect about HIV-risk, then their perceptions about age-similar versus age-disparate partners may also be flawed. Successful public health interventions to meet the UNAIDS 90-90-90 targets, which aim to have 90% of HIV positive people diagnosed by 2020 (230,231), will depend on the public's ability to accurately determine which sexual behaviours put them at risk, since these risk-perceptions also influence health-seeking behaviours. Though addressing HIV-risk perceptions through public educational campaigns is necessary, it may not be sufficient (See recommendation 3, below).

3. Informational campaigns that raise awareness about context-specific HIV behavioural risk factors, should not assume a moralizing tone, nor should they be based on the premise that women have low self-esteem (88). If anything, qualitative investigations have demonstrated that in many places women exercise agency in their choice of older partners, and are often intelligently balancing various risks with which they are presented in different circumstances (14,88,232). As I have seen first-hand, Cape Town women living in disadvantaged communities may perceive an older man who does not want to use a condom and who has concurrent relationships as less risky, in terms of bad relationships, than an age-similar man who might abuse her. In the first instance, the risk is far-off and HIV transmission may never occur, while the second instance may seem to be more immediate and direct.
4. There is a need to strengthen cultural, economic, educational and legal structures to protect and empower women (97). While there is evidence to suggest women have control over partnership formation, as well as the number and type of partners they have, within these relationships they have less power in condom use negotiations and sex frequencies (14). Young women in Southern Africa have growing aspirations, and limited opportunities for financial independence. Programmes incentivizing women to stay in school longer, like state-sponsored cash-transfers, or start businesses, like micro-lending programmes, are examples of some avenues that may lead to financial independence, and thus, increased relationship power. However, if these programmes are not sustained over the long-term, so that young women have the opportunity to develop useful skills for a modern economy or acquire productive assets, then they may not be effective.
5. There is also a need for interventions in Cape Town that address psychological consequences and social determinants of intimate partner violence, as well as provide legal solutions for women who are victims. Though we do not know if age similar partners are more abusive than age-disparate partners, the perceptions of young women guide their actions, and point to their real fears about intimate partner violence. On top of this, it has been suggested that sexual risk reduction strategies may place women at further risk for IPV if they start making specific demands of their partners (233). Again, this points to the need for complex interventions that provide help to abused women, while simultaneously setting women up for financial independence, and improved ability to negotiate safer sex within relationships. At the same time, interventions are also needed to curb the perpetration of abuse by men. Interventions starting at young ages, well before men begin intimate relationships, and which aim to prevent antisocial patterns are needed. Additionally, interventions for adult men that help them to deal with anger and also increase their parenting skills are warranted in this context (234).
6. Thus far most of the intervention suggestions and framings of this topic in the literature have been gynocentric. However, our studies also point to the need to protect young men from HIV transmission by older or age-similar partners. One of the most efficacious interventions for preventing transmission from women to men is medical male circumcision (235–237). Voluntary medical male circumcision programmes are currently underway, but there are still several

barriers to scale-up. Lack of funding, programme inefficiencies, demand creation, low capacity of government to manage programmes, as well as time consuming and invasive medical procedures are all seen as shortcomings of the current programmes (238).

These intervention strategies have the potential to curb HIV incidence and potentially eliminate key transmission pathways in certain age groups, but they will require a great deal of political will to accomplish them.

## 9.2 CONCLUSIONS

---

Age differences in heterosexual relationships have the potential to both, elevate the individual-level risk of acquiring HIV, and help sustain the epidemic within some populations of South Africa and Malawi. Although, the way in which age differences serve these ends may be different than what has been previously described in observational studies. We now recognize that all partner ages – not only the most recent or primary partner – are important, and crucially that partner age variability is probably a key driver of sustained epidemics. Moreover, we know in some of the contexts explored here, women may be seeking out relationships with older men because of the financial and emotional security they are provided. Those relationships are often long-term and may involve many sexual encounters without condoms. When those women become infected with HIV, they may then transmit the virus to men in age-similar relationships, thus perpetuating the cycle of transmission.

The high HIV incidence among young women is a multi-faceted problem, that cannot be addressed whole-sale with campaigns to stigmatize sugar daddy relationships. Understanding age-mixing dynamics in high HIV prevalence settings is the key to identifying transmission pathways in those populations. Awareness of the male and female motives for partner age variability and preferences should be the basis for interventions eliminate those transmission chains.

## ABSTRACT

---

Southern Africa has an immense HIV epidemic marked by a disproportionate amount of new infections among adolescent and young adult women. The high prevalence of women engaging in age-disparate relationships – where men are five or more years older than women – has been suggested as a cause of young women acquiring HIV. Age-disparate relationships have been linked to inconsistent or less condom use in the relationship, multiple partner concurrency by the older male partner, higher sex frequencies, and alcohol use before sex. Moreover, in South Africa and Malawi HIV prevalence peaks in older age groups for men compared to women, implying that women who choose an older partner will have a greater likelihood of choosing an HIV infected partner. Compounding all of this, is the fact that younger women may be more susceptible to HIV infection than men and older women because they tend to have a different mucosal immune environment in their genitals and they are also prone to genital micro-abrasions. Despite these theorized mechanisms for how age-disparate relationships may cause HIV infection in young women, the epidemiological evidence has been inconclusive, with the most recent and best evidence suggesting there is no relationship between age differences and incident HIV infection among young women.

There are several key limitations of previous work, that necessitated a closer investigation and new approach to studying age differences and HIV infection. Previous observational studies have not provided detailed descriptions of age-mixing patterns – population-level patterns for how people choose partners with regards to age – which may help to explain the magnitude and persistence of HIV in a population. Additionally, most studies have had serious methodological shortcomings, such as using face-to-face interviewing or only examining risk of HIV infection in the most recent sexual relationships. The former limitation may result in social desirability bias in reporting sexual behaviours, and the latter does not consider the risk of HIV posed by previous non-recent, age-disparate relationships. Finally, previous epidemiological studies have neglected to consider the agency of women in these relationships, and ultimately the role they might play in sustaining the HIV epidemic. To this end, this thesis explored new data from different South African and Malawian populations to evaluate whether age differences in sexual relationships have the potential to maintain the HIV epidemic in Southern Africa and augment the individual-level risk of acquiring HIV.

I conducted one qualitative study consisting of in-depth interviews with key informants from Cape Town, as well as four secondary statistical analyses of data and from: Cape Town, South Africa (CAPS and CTSBS studies); Zomba, Malawi (SIHR study); and Likoma Island, Malawi (LNS study). Using this quantitative and qualitative data I: 1. Described and visualized age-mixing patterns in the CAPS, LNS, and CTSBS populations (Chapters 3, 5 and 7); 2. Identified socio-demographic predictors and psycho-social motivations for engaging in relationships with older men in the CAPS and CTSBS settings (Chapters 3 and 4); 3. Ascertained whether age differences in relationships were correlated with an individual's HIV infection status in the LNS and CTSBS studies (Chapters 5 and 7), or other sexual risk behaviours commonly associated with increased HIV transmission risk in the CAPS, LNS, and SIHR populations (Chapters 3, 5 and 6). Where possible I attempted to address the limitations of previous studies by using data collected with state-of-the-art interviewing techniques, such as

Audio Computer-Assisted Self-Interviewing, as well as advanced statistical techniques, like multivariate imputation by chained equations, generalized additive models, and mixed effects models.

We found that large age differences in relationships were associated with less condom use in relationships (Chapters 3, 5, and 6), more frequent sex (Chapter 6), longer relationship durations (Chapter 6), and more reported STI symptoms (Chapter 3). Additionally, men who had a large range of age differences in the LNS population, were more likely to report having a concurrent relationship in the previous 3 years (Chapter 5). These sexual risk behaviours may put individuals at increased risk of HIV infection when they are in relationships with older partners. Indeed, in the two studies for which we had access to the HIV status of participants, we found a relationship between age differences and HIV infection. In the LNS study we observed that men, aged 18-49, who had relationships with older women were more likely to be HIV positive (Chapter 5). For women, aged 18-49, they had increasing probabilities of being HIV infected as the age differences grew from 2 to 12 years, but then they had decreasing probabilities of HIV infection as the age differences grew larger. In the CTSBS population, women aged 15-70 who were HIV positive had nearly three times the expected range of partner ages, as those who were HIV negative (Chapter 7). At the population level, we observed that there was larger within-subject variation in age differences for women compared to men in the LNS population (Chapter 5), and larger within-subject variation in age differences for HIV positive women compared to HIV negative women in the CTSBS setting (Chapter 7). The large variation in age differences for women may be an important characteristic of large, sustained HIV epidemics in Southern Africa.

Our study of the two Cape Town populations indicates that precarious and unstable home lives of young women, possibly characterized by absent parents, destitution, food insecurity, and intimate partner violence may be key determinants of women seeking relationships with older men (Chapters 3 and 4). Providing these young women with other forms of income through cash-transfers or state-sponsored grants may bring increased stability to young women, and eliminate the some of the motivation for relationships with older, riskier men who are likely to infect them (Chapter 6). By eliminating this transmission chain, there may be trickle-down effects to younger men, who will also be less likely to become HIV infected from their age-similar, or older, female partners.

Cash transfers are just one type of potential interventions that has the ability to prevent age-dependent chains of transmission. My research suggests other possible guidelines for future HIV prevention interventions related to age differences in relationships. Interventions should: 1. Be context dependent and evidence-based as age-mixing patterns differ across settings; 2. Address inaccuracies in HIV- and relationship-risk perceptions; 3. Not be moralizing or assume young women have no agency in partner selection; 4. Strengthen legal and economic structures that empower women to make safer sexual behaviour choices within relationships; 5. Address young women's concerns about intimate partner violence; and 6. Consider the HIV risk posed to young men in age-similar relationships. Policy recommendations should be enhanced by evidence from relationship data that have been derived from well-conducted sexual behaviour surveys, and analyses utilizing appropriate statistical techniques for handling missing data and modelling non-linear associations.

## SAMENVATTING

---

Zuidelijk Afrika heeft een enorm grote hiv-epidemie, gekenmerkt door een disproportioneel aantal nieuwe infecties onder jonge vrouwen. Het frequent voorkomen van vrouwen die relaties aangaan met mannen die vijf of meer jaar ouder zijn dan zijzelf, werd reeds aangehaald als mogelijke oorzaak voor het hoge risico op hiv-infectie bij jonge vrouwen. In deze zogenaamde “relaties met ongelijke leeftijd” (ROLs) komt inconsistent of verminderd condoomgebruik vaker voor. Andere factoren geassocieerd aan ROLs zijn parallele relaties aangegaan door de oudere mannelijke partner, een hogere frequentie van seksuele gemeenschap en het gebruik van alcohol vóór seks. Daarnaast is het zo dat in Zuid-Afrika en Malawi de prevalentie van hiv piekt op latere leeftijd bij mannen dan bij vrouwen. Daardoor hebben vrouwen die een oudere partner kiezen een hogere kans dat deze partner hiv-geïnfecteerd is. Bovendien is het zo dat jonge vrouwen mogelijk vatbaarder zijn voor hiv-infectie dan mannen en oudere vrouwen, ten gevolge van verschillen in het afweersysteem in de slijmvliezen van hun genitaliën, en hun geneigdheid tot het oplopen van microscopisch kleine schaafwonden ter hoogte van de genitaliën. Ondanks deze beredeneerde mechanismen voor de manier waarop ROLs hiv-infecties kunnen veroorzaken bij jonge vrouwen, is epidemiologisch bewijsmateriaal niet conclusief, met de meest recente en beste aanwijzingen dat er geen relatie bestaat tussen de grootte van de leeftijdsverschillen tussen partners enerzijds, en de incidentie van hiv-infecties bij de vrouwelijke partners anderzijds.

Voorafgaand werk vertoont verscheidene beperkingen die een dieper onderzoek en een nieuwe aanpak noodzakelijk maken voor het bestuderen van de link tussen leeftijdsverschillen en hiv-infectie. Eerdere observationele studies bevatten geen gedetailleerde beschrijving van het patroon van leeftijdsverschillen tussen partners. Dit patroon verklaart mogelijk deels de grootte en persistentie van hiv-epidemieën. Daarnaast hebben de meeste studies belangrijke methodologische tekortkomingen, zoals het gebruik van “face-to-face” interviews of enkel het beschouwen van de meest recent seksuele relatie bij het bestuderen van het risico op hiv-infectie. De eerste beperking kan leiden tot vertekening van zelf-gerapporteerd seksueel gedrag, naar sociaal-wenselijke antwoorden. De tweede beperking houdt geen rekening met het risico op hiv-infectie door voorgaande minder recente relaties. Tot slot hebben eerdere epidemiologische studies de actieve rol van vrouwen bij het maken van beslissingen rond relaties verwaarloosd, en dus de rol die ze mogelijk spelen bij het in stand houden van hiv-epidemieën. Deze thesis exploreert nieuwe gegevens van bevolkingsgroepen uit Zuid-Afrika en Malawi, en evalueert de hypothese dat leeftijdsverschillen tussen sekspartners een rol spelen bij het in stand houden van de hiv-epidemieën in Zuidelijk Afrika en dat ze het risico op hiv-infectie verhogen.

Ik heb één kwalitatieve studie uitgevoerd, bestaande uit diepte-interviews met informanten uit Kaapstad (Zuid-Afrika), evenals vier secundaire statistische analyses van gegevens uit Kaapstad (CAPS en CTSBS studies); Zomba (Malawi: de SIHR studie), en het eiland Likoma (Malawi: de LNS studie). Op basis van deze kwalitatieve en kwantitatieve gegevens heb ik: (1) de patronen van leeftijdsverschillen tussen partners beschreven en gevisualiseerd in de CAPS, LNS en CTSBS studies (hoofdstukken 3, 5 en 7); (2) socio-demografische predictoren en psychosociale motivaties geïdentificeerd voor het aangaan van relaties met oudere mannen in de CAPS en CTSBS studies

(hoofdstukken 3 en 4); (3) onderzocht of leeftijdsverschillen tussen sekspartners gecorreleerd zijn aan iemands hiv-infectie status in de LNS en CTSBS studies (hoofdstukken 5 en 7), of aan seksueel risicogedrag dat een verhoogde kans op hiv-overdracht met zich meebrengt in de CAPS, LNS en SIHR studies (hoofdstukken 3, 5 en 6). Waar mogelijk heb ik geprobeerd om de beperkingen van voorgaande studies te overstijgen, door gegevens te gebruiken die verzameld werden via audio-computergeassisteerde interviews, en door geavanceerde statistische methoden te gebruiken, zoals multivariate imputatie door gekoppelde vergelijkingen, gegeneraliseerde additieve modellen, en gemengde-effecten modellen.

Volgens onze bevindingen waren grotere leeftijdsverschillen tussen sekspartners geassocieerd met minder condoomgebruik (hoofdstukken 3, 5 en 6), hogere frequentie van seks (hoofdstuk 6), meer langdurige relaties (hoofdstuk 6) en meer zelf-gerapporteerde symptomen van seksueel overdraagbare infecties (hoofdstuk 3). Bovendien hadden mannen die een grote spreiding aan leeftijdsverschillen met hun partners hadden, meer kans om betrokken te zijn geweest in parallelle relaties in de voorbije drie jaar (hoofdstuk 5). Deze seksuele risicogedragingen verhogen mogelijks de kans op hiv-infectie voor vrouwen die relaties aangaan met oudere mannen. In de twee studies waarvoor we toegang hadden tot de gegevens omtrent de hiv-status van de deelnemers vonden we inderdaad een associatie tussen grotere leeftijdsverschillen en de kans op hiv-infectie. In de LNS studie observeerden we dat mannen die relaties waren aangegaan met oudere vrouwen meer kans hadden om hiv-positief te hebben getest (hoofdstuk 5). Voor vrouwen uit de LNS studie nam de kans op hiv-infectie ook toe naarmate hun leeftijdsverschillen met sekspartners toenamen van twee tot twaalf jaar, maar voor nóg grote leeftijdsverschillen nam de kans op hiv-infectie weer af.

Hiv-geïnfecteerde vrouwen in de CTSBS studie hadden gemiddeld een spreiding van leeftijdsverschillen met partners die bijna drie keer zo groot was als het gemiddelde voor hiv-negatieve vrouwen (hoofdstuk 7). Op populatieniveau was er een grotere individuspecifieke variatie in leeftijdsverschillen bij vrouwen dan bij mannen in de LNS studie (hoofdstuk 5), en een grotere individuspecifieke variatie in leeftijdsverschillen bij hiv-geïnfecteerde vrouwen in vergelijking met hiv-negatieve vrouwen in de CTSBS studie (hoofdstuk 7). Deze grote variatie in leeftijdsverschillen bij vrouwen is mogelijks een belangrijk kenmerk van grote, langdurige hiv-epidemieën in Zuidelijk Afrika.

Onze twee studies in Kaapstad geven aan dat ongunstige, onstabiele thuissituaties, gekenmerkt door afwezige ouders, armoede, voedselonzekerheid, intiem partnergeweld, sleutelfactoren kunnen zijn waarom jonge vrouwen relaties zoeken met oudere mannen (hoofdstukken 3 en 4). Het verstrekken van andere bronnen van inkomsten, via contant betalingen of subsidies van de overheid, zou verhoogde stabiliteit kunnen betekenen voor deze jonge vrouwen, en zou dus sommige van de motiverende factoren voor het aangaan van meer riskante relaties met oudere mannen (want een verhoogde kans op hiv-overdracht) kunnen elimineren (hoofdstuk 6). Het elimineren van de transmissieketen kan mogelijks ook neveneffecten hebben voor jongere mannen, die dan minder kans zouden hebben op hiv-besmetting.

Contant betalingen zijn slechts één soort van potentiële interventies die de mogelijkheid hebben om leeftijdsafhankelijke hiv-transmissieketens te doorbreken. Mijn onderzoek suggereertzets mogelijke richtlijnen voor toekomstige interventies ter preventie van hiv infecties, gericht op leeftijdsverschillen tussen sekspartners. Interventies moeten: 1. Context-afhankelijk en gebaseerd zijn op objectieve gegevens, omdat patronen van

leeftijdsverschillen verschilt tussen omgevingen; 2. Onjuiste inschattingen en percepties van het risico op hiv-overdracht en andere relatie-gebonden risico's corrigeren; 3. Niet moraliserend zijn en niet aannemen dat vrouwen geen actieve rol spelen bij het kiezen van partners; 4. Legale en economische structuren versterken die vrouwen in staat stellen om veiligere keuzes te maken omtrent seksueel gedrag binnen relaties; 5. Werk maken van de bezorgdheden van jonge vrouwen inzake intiem partnergeweld; en 6. Rekening houden met het risico op hiv-overdracht naar jonge mannen in relaties met leeftijdsgenoten. Beleidsaanbevelingen dienen ondersteund te zijn door wetenschappelijke onderzoek met behulp van goed-uitgevoerde bevestigingen omtrent seksueel gedrag, en statistische methodes die op gepaste wijze rekening houden met ontbrekende gegevens en niet-lineaire associaties tussen variabelen.

## ACKNOWLEDGEMENTS

---



Photo credit: IFPRI-IMAGES via VisualHunt.com / CC BY-NC-ND

This dissertation took a tremendous amount of persistence, learning, and effort over the past four years and I would not have been able to finish this work if not for the assistance, advice and support from numerous people along the way. First and foremost, I would like to thank my Promoter, Prof. Wim Delva. More than any other single person, he has done the most to get me to this point in my career. He has gone above and beyond the call of duty as my supervisor. In addition to providing expertise and support with research activities, he has encouraged and advocated for my career and professional growth, made sure I have had stable employment, income, and a good work-life balance, as well as been a friend and confidant. I suspect this is not where our relationship will end, and I hope we have many years of fruitful collaboration ahead of us.

I also would like to extend my gratitude to my co-promotor, Prof. Niel Hens. Niel is a renowned statistician who has been instrumental to the development of my statistical skills. Though I am not a classically trained statistician, he has never looked down on me (except in the literal sense, since legend has it, he is related to giants), nor made me feel foolish for asking too many questions or analysing data incorrectly. No matter how busy he is, or how high he rises through the ranks in his own career, he is someone I could always depend on to respond to my emails in a timely manner, answer my questions and give advice on all of my statistical quandaries.

I would also like to express my sincerest appreciation for the International Centre for Reproductive Health (ICRH) and my other co-promoter, Prof. Marleen Temmerman. I have to give credit to Marleen for building such a preeminent research institution, that has been my “nourishing mother” over the past four years. Marleen has helped to make it possible for me to conduct the research I do by extending her network to myself, and thereby helping me to secure funding and collaborations with distinguished researchers at other institutions. Despite the fact that I have been residing in South Africa during my PhD, the ICRH has always made me feel included and part of their community. I especially have to thank Dirk Van Braeckel, Cindy De Muynck, and Chris Moreel at ICRH for countless hours of PhD programming and administrative support.

I would also like to express my deepest gratitude to the South African Centre for Epidemiological Modelling & Analysis (SACEMA). SACEMA has been my work home for the past seven years and has provided me with unending professional support and capacity development. I will never forget where my career truly started. Prof. Delva and Prof. Alex Welte took a chance hiring me when I had not yet completed my MPH, and they did their best to make sure SACEMA was always a pleasant, friendly, and encouraging environment to work in. I will forever be grateful for the lessons learned at SACEMA in how to collaborate with others; teach and supervise students; and be critical of data and models. Others at SACEMA also deserve special thanks: Prof. John Hargrove, Dr. Gavin Hitchcock, Cari van Schalkwyk, Amanda October, and Lynnemore Scheepers. They have been with me since the beginning of my SACEMA adventure and they have become my friends, colleagues, mentors, and co-commiserators.

Through SACEMA I was also introduced to the International Clinics on Infectious Disease Dynamics and Data (ICI3D) programme. I would like to thank the outstanding and distinguished faculty associated with that programme: Prof. Juliet Pulliam for running the programme and providing me with several opportunities over the years to learn about dynamical modelling; Prof. Jonathan Dushoff for hosting me during an exchange at McMaster university and teaching me so much about statistical modelling;

Dr. Brian Williams and Prof. Steve Bellan for their excellent mentorship in epidemiological modelling and analysis; and finally, Prof. Travis Porco for a wonderful and unexpected collaboration resulting in a manuscript.

I would also like to thank my life partner, Matthys Kroon, for always providing emotional support and encouragement throughout my work and studies. Though Matthys is not doing work in the same field as me, his ability to understand and engage with ideas on fundamental and deep level has meant that I always had a willing person to bounce ideas off of or someone to engage in critical discussions with about ethics and research. I would also be remiss, if I did not mention how often he has had to take over household responsibilities, including taking care of my fur children, while I was away travelling for work. His willingness to do this for me has allowed me to take advantage of many networking and capacity development opportunities that have come my way.

Finally, I give special thanks to my family: Susan James (mother), Stephen Beauclair (father), Veronica and David Saxberg (grandparents), Anna Beauclair (grandmother), DesiRae Beauclair (sister), Jade Beauclair (sister), Penny and Joel Neumiller (aunt and uncle) for the values they have instilled and fostered within me that I truly believe are what allowed me to conquer this giant I call my dissertation. In their own ways, they have influenced my love of reading and learning; my ability to think outside of the box and be critical of ideas and institutions, as well as the status quo.

*Gratias maximus vobis ago!*

## LIST OF PUBLICATIONS

---

1. **Beauclair, R.**; Helleringer, S.; Hens, N.; Delva, W. (2016) 'Age differences between sexual partners, behavioural and demographic correlates, and HIV infection on Likoma Island, Malawi'. *Scientific Reports*: 6. (IF: 5.47)
2. Aboyade, O.M; **Beauclair, R.N**; Mbamalu, O.N; Puoane, TR; Hughes, G.D. (2016) 'Health-seeking behaviours of older black women living with non-communicable diseases in an urban township in South Africa.' *BMC Complementary and Alternative Medicine*; 16(410):1-9.
3. Brook, C.; **Beauclair, R.**; Ngwenya, O.; Worden, L.; Ndeffo-Mbah, M.; Lietman, T.; Satpathy, S.; Galvani, A.; and Porco, T. (2015) 'Spatial heterogeneity in projected leprosy trends in India'. *Parasites & Vectors*: 8(1):1-11. (IF: 3.43)
4. Hughes, G.; Aboyade, O.M.; **Beauclair, R.**; Mbamalu, O.N.; Puoane, T.R. (2015) 'Characterizing Herbal Medicine Use for Noncommunicable Diseases in Urban South Africa'. *Evidence-Based Complementary and Alternative Medicine*: 1-10 (IF: 1.88)
5. **Beauclair, R.**; Hens, N.; Delva, D. (2015) 'Concurrent partnerships in Cape Town, South Africa: race and sex differences in prevalence and duration of overlap' (2015). *Journal of the International AIDS Society* 18(1): 19372. (IF: 5.10)
6. **Beauclair, R.**; Petro, G.; Myer, L. (2014) 'The association between timing of initiation of antenatal care and stillbirths: A retrospective cohort study of pregnant women in Cape Town, South Africa'. *BMC Pregnancy and Childbirth* 14:204. (IF: 3.41)
7. **Beauclair, R.** and Delva, W. (2013) 'Is younger really safer? A qualitative study of perceived risks and benefits of age-disparate relationships among women in Cape Town, South Africa'. *PLoS ONE* 8(11): e81748. (IF: 3.20)
8. Delva, W.; Meng, F.; **Beauclair, R.**; Deprez, N.; Temmerman, M.; Welte, A.; Hens, N. (2013) 'Coital frequency and condom use in monogamous and concurrent sexual relationships in Cape Town, South Africa.' *Journal of the International AIDS Society* 16: 18034. (IF: 5.10)
9. **Beauclair, R.**; Meng, F.; Deprez, N.; Temmerman, M.; Welte, A.; Hens, N.; Delva, W. (2013) 'Evaluating Audio Computer Assisted Self-Interviews in Urban South African Communities: Evidence for Good Suitability and Reduced Social Desirability Bias of a Cross-sectional Survey on Sexual Behaviour.' *BMC Medical Research Methodology* 13:11. (IF: 2.67)
10. Michielsen, K.; **Beauclair, R.**; Delva, W.; Van Rossem, R.; Temmerman, M. (2012) 'Effectiveness of a peer-led HIV prevention intervention in secondary schools in Rwanda: results from a non-randomized controlled trial.' *BMC Public Health* 12:729. (IF: 2.00)
11. **Beauclair, R.**; Kassinjee, R.; Welte, A.; Temmerman, M.; Delva, W. (2012). 'Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa.' *The European Journal of Contraception and Reproductive Health Care* 17(1):30. (IF: 1.46)
12. Delva, W.; **Beauclair, R.**; Welte, A.; Vansteelandt, N.; Hens, N.; Aerts, M.; du Toit, E.; Beyers, N.; Temmerman, M. (2011). 'Age disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol.' *BMC Public Health* 2(11): 616. (IF: 2.00)

## RESEARCH ARTICLE

## Open Access

## Evaluating audio computer assisted self-interviews in urban south African communities: evidence for good suitability and reduced social desirability bias of a cross-sectional survey on sexual behaviour

Roxanne Beauchair<sup>1\*</sup>, Fei Meng<sup>2,3</sup>, Nele Deprez<sup>4</sup>, Marleen Temmerman<sup>2</sup>, Alex Welte<sup>1</sup>, Niel Hens<sup>3,5</sup> and Wim Delva<sup>1,2</sup>

### Abstract

**Background:** Efficient HIV prevention requires accurate identification of individuals with risky sexual behaviour. However, self-reported data from sexual behaviour surveys are prone to social desirability bias (SDB). Audio Computer-Assisted Self-Interviewing (ACASI) has been suggested as an alternative to face-to-face interviewing (FTFI), because it may promote interview privacy and reduce SDB. However, little is known about the suitability and accuracy of ACASI in urban communities with high HIV prevalence in South Africa. To test this, we conducted a sexual behaviour survey in Cape Town, South Africa, using ACASI methods.

**Methods:** Participants (n = 878) answered questions about their sexual relationships on a touch screen computer in a private mobile office. We included questions at the end of the ACASI survey that were used to assess participants' perceived ease of use, privacy, and truthfulness. Univariate logistic regression models, supported by multivariate models, were applied to identify groups of people who had adverse interviewing experiences. Further, we constructed male–female ratios of self-reported sexual behaviours as indicators of SDB. We used these indicators to compare SDB in our survey and in recent FTFI-based Demographic and Health Surveys (DHSs) from Lesotho, Swaziland, and Zimbabwe.

**Results:** Most participants found our methods easy to use (85.9%), perceived privacy (96.3%) and preferred ACASI to other modes of inquiry (82.5%) when reporting on sexual behaviours. Unemployed participants and those in the 40–70 year old age group were the least likely to find our methods easy to use (OR 0.69; 95% CI: 0.47–1.01 and OR 0.37; 95% CI: 0.23–0.58, respectively). In our survey, the male–female ratio for reporting >2 sexual partners in the past year, a concurrent relationship in the past year, and > 2 sexual partners in a lifetime was 3.4, 2.6, and 1.2, respectively— far lower than the ratios observed in the Demographic and Health Surveys.

**Conclusions:** Our analysis suggests that most participants in our survey found the ACASI modality to be acceptable, private, and user-friendly. Moreover, our results indicate lower SDB than in FTFI techniques. Targeting older and unemployed participants for ACASI training prior to taking the survey may help to improve their perception of ease and privacy.

**Keywords:** ACASI, Sexual behaviour, Social desirability bias, Self-reported data, Gender, South Africa

\* Correspondence: [roxanne.beauchair@gmail.com](mailto:roxanne.beauchair@gmail.com)

<sup>1</sup>The South African Department of Science and Technology/National Research Foundation (DST/NRF) Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, 19 Jonkershoek Road, Stellenbosch 7600, South Africa

Full list of author information is available at the end of the article



© 2013 Beauchair et al.; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Background

Sexual behavioural research largely depends on self-reported data in the absence of more objective biomarkers. Ensuring the validity of these data is a common challenge faced by traditional survey methods. Specifically, self-reports of sensitive issues such as engagement in multiple, concurrent sexual relationships may be subject to social desirability bias (SDB)—or the inclination to report behaviours that are viewed in a favourable way by society. For sexual behaviours the bias is often gendered because of sexual double standards in society, which frequently lead men to over-report and women to under-report certain behaviours [1-4].

For several countries in Africa, young unmarried women tend to report a later age for their first sexual experience compared to young unmarried men [5]. In regard to sexual concurrency—overlapping sexual relationships where a new sexual relationship starts before an existing relationship ends—several studies from Africa indicate that men are far more likely to report having multiple sexual partners at once, or extra-marital partners, compared to women [6-9]. Women have also been found to report fewer non-spousal partners in the past year compared to men [10]. Gendered discrepancies also exist in reporting of condom use in Africa. Men consistently are more likely to report having ever used condoms [6] and always using condoms with all of their partners [9,11] than women. A study conducted specifically to test SDB among married couples, found that a higher percentage of husbands reported that they used condoms with their spouse, than their wives did [12].

When participants consciously modify their answers to project a socially 'appropriate' response, this is called impression management [13], and the effects of it can theoretically be ameliorated by changing the design of the study. To combat the problem of SDB, several studies have investigated how mode of inquiry might reduce this phenomenon. Audio Computer-Assisted Self-Interviewing (ACASI) techniques have emerged as a way to curtail SDB because they allow participants to answer questions in privacy by using a computer to read the question, while simultaneously listening to an audio recording of the questions being asked. Several studies have attempted to evaluate the accuracy of ACASI-generated survey results and have produced disparate results. For example, studies done in India have shown that women reported forcibly being touched [14] and domestic violence less often in ACASI compared to traditional Face-to-face-interviewing (FTFI) methods [15]. In the U.S. a study was conducted to see if reported condom use in an ACASI could be substantiated by a Y-chromosome polymerase chain reaction (Yc-PCR) test to detect sperm in vaginal fluid. It demonstrated that even with ACASI methods, there were still inconsistencies between the

Yc-PCR and reported results [16]. However, in the African context ACASI methods seemed to precipitate the reporting of more sexual behaviours [17]. Increases in the reporting of number of sexual partners [18], two or more sex partners at once [19], and for women, having had sex with a relative, stranger, or older man [20] in ACASI compared to FTFI seem to indicate that participants in Africa may be more comfortable answering questions about sexual behaviour when they have access to greater privacy.

These discrepant findings about ACASI compared to FTFI are indicative of the need for more context-specific research on ACASI methods. Particularly, since little is known about the validity of ACASI and its suitability for the context of urban, disadvantaged communities in South Africa, where the burden of HIV/AIDS is among the highest in the world. We conducted a cross-sectional survey in three urban communities of Cape Town, South Africa that are known to have a high prevalence of HIV. As a part of the survey, we asked participants questions about their interview experience. In this paper, we report the results of these questions and we compare key summary statistics known to be highly sensitive to SDB from our survey with those from the Demographic and Health Surveys (DHS), conducted in Lesotho, Swaziland, and Zimbabwe [21-23].

## Methods

### The Cape Town sexual behaviour ACASI survey

We administered a cross-sectional sexual behaviour survey ( $n = 878$ ) in three urban, economically disadvantaged communities around Cape Town from June 2011 to February 2012. The survey took a sub-sample of participants from the Zambia South Africa TB and AIDS Reduction Study (ZAMSTAR), a community randomised trial that aimed to reduce the prevalence of TB and HIV using novel public health interventions. The primary objective of the survey analyzed here was to look at associations between HIV status, sexual connectedness and age-disparate relationships. To this end, we explored participants' one-year sexual histories focusing on dates of sexual activity episodes, coital frequency, condom use, age differences between partners, as well as drug and alcohol use at the time of first sexual intercourse. The complete study design and protocol are explained in further detail elsewhere [24].

The questionnaire was conducted in a safe, mobile office space using ACASI on touch screen computers. The ACASI software was developed by an independent software engineering company. Participants listened to audio recordings of questions and response options through headphones. They were allowed to take the survey in their choice of language: English, Afrikaans, or isiXhosa. Furthermore, the questionnaire was set up along a temporal trajectory with the onset, dissolution,

and duration of each relationship episode displayed on a touch screen timeline for the participant to see, using different colours for each partner. Participants had the option of listening to the questions over again, or skipping the question if they did not want to answer it. At the end of the survey we added four questions to assess the participant's experience with the ACASI mode of inquiry. It has been established that any survey method needs to be user-friendly and offer privacy, if participants are to answer truthfully and give full attention to the questions for the entire duration of the questionnaire [13]. Therefore, we surmised that these four questions would give the best indication of whether or not participants had a positive experience with a survey modality that was new to their community.

Of 1857 people randomly sampled from the ZAM-STAR sampling frame, we were able to locate the addresses of 1115 (60.0% contact rate). For 197 people, the reason for non-contact after three attempts was unknown, while for respectively 511 and 34, relocation to an unknown new address and death was documented. An additional 87 candidate participants were excluded, primarily due to visual or physical impairments that rendered participation in the study impossible. Of the remaining 1028 who were eligible, 878 (85.4% response rate) took the survey [25].

For this analysis, we excluded participants that had missing ages ( $n = 7$ ), were younger than 15 years old ( $n = 1$ ), and were older than 70 years old ( $n = 26$ ). We also excluded participants that did not identify as a man or woman ( $n = 30$ ). Additionally, we excluded participants ( $n = 16$ ) that did not identify as a black or coloured person, as these racial groups are not typical of populations that have a high prevalence of HIV in South African urban communities. The term *coloured* refers to racially mixed descendants of Europeans, indigenous populations and slaves from South and East Asia.

#### Statistical analysis

Our analysis is divided into three parts. First we computed descriptive statistics for the four questions that assessed the participants' interview experience. We asked about the participant's ease of use (very easy/somewhat easy/somewhat difficult/very difficult/I don't know), perceived confidentiality (very private/somewhat private/no privacy/I don't know) and self-reported truthfulness of their answers to the survey questions (all questions answered truthfully/most truthfully/some truthfully, some dishonestly/most dishonestly/all dishonestly/I don't know). Finally, we asked them to indicate their preferred medium of answering questions about sexual behaviour (touch screen computer with ACASI/researcher-administered verbal questionnaire/self-administered written questionnaire/telephonic survey/I don't know).

Second, we conducted logistic regression analyses to identify groups of people that were less likely to have a positive experience with the interview. Five different univariate models were constructed for each of the following outcome variables: ease of use (easy/difficult), perceived confidentiality (private/no privacy), truthfulness (all truthfully/some dishonesty), and mode of inquiry (ACASI/other type). The univariate models examined the following variables as predictors of each of the outcomes: race (black/coloured), age (15–24/25–39/40–70), gender (male/female), education level (primary/secondary/tertiary), and employment status (employed/unemployed). Next, we constructed multivariate logistic regression models, adjusting for all of the predictors. However, here we present the results of the univariate analysis because we were primarily concerned with finding predictors of adverse experiences using marginal associations. Additionally, we thought that unadjusted ORs would be easier for others to interpret for their own settings if they are looking for 'at-risk' individuals to target for training in their own ACASI survey.

Finally, we attempted to assess the extent to which our ACASI survey was successful in reducing SDB. While it is impossible to ascertain how accurate our results were, since the truth is unknown, we can attempt relative assessments of accuracy by comparing responses, by gender, among results from surveys that use different modes of inquiry. To this end, we constructed male–female prevalence ratios of self-reported sexual behaviours, as indicators of SDB, and we used these indicators to compare SDB in our survey and in recent FTFI-based, nationally representative DHSs from Lesotho, Swaziland, and Zimbabwe [21–23]. These three countries were chosen as comparison cases based on the assumptions that their cultural and behavioural context was most similar to South Africa. Specifically, we compared the male–female ratios for the variables: *> 2 partners in the past year*, *> 2 total lifetime number of sexual partners*, and *had a concurrent relationship in the past year*. Corresponding with the DHS data analysis, described in detail elsewhere [26], we defined concurrency as having two or more partners that overlapped in time in the year preceding the survey. The ratios were constructed by dividing the prevalence of the behavior reported for men, by the prevalence of the behavior reported for women. For this section of the analysis we limited our study population to the same categories of participants included in the DHS working paper on concurrent sexual partnerships in order to make our results more directly comparable with theirs. Thus, we made additional exclusions of participants who were coloured ( $n = 196$ ) and older than 49 years old ( $n = 99$ ).

For each of these variables, we assumed that men tend to over-report sexual behaviours, and women tend to under-report them. If the prevalence's of behaviours

for men and women were in truth not as divergent as reported by FTFI-based surveys, we would expect these surveys to overestimate the true male–female ratios. Therefore, if our survey, that used ACASI methods, produces smaller male–female ratios for these behaviours, compared to the DHS surveys, we could expect our estimate to be closer to the truth, since male over-reporting and female under-reporting is minimized. We will calculate 95% confidence intervals for all ratios using Fisher's exact methods of estimation. All data analyses were performed using R [27].

#### Ethical approval

The study was approved by the Stellenbosch University Health Research Ethics Committee (N11/03/093). Written, informed consent was obtained for each respondent prior to administration of the questionnaire.

#### Results

After exclusions, we were able to evaluate the ACASI user experience for 798 participants. Our study population was comprised mostly of women (68%) and participants identifying as black (75%). The median age of participants was 35.5 years old (IQR 26–46). Most participants chose to take the survey in isiXhosa language (66.5%), followed by Afrikaans (21.9%) and English (11.4%). Table 1 displays the results from our questions about interview modality. A majority of participants found that using ACASI was easy ('very easy': 67.2% and 'somewhat easy': 18.7%). Most participants also found that the ACASI method offered a certain level of privacy ('very private': 88.7% and 'somewhat private': 7.6%). Eighty-six percent of the participants claimed to have answered all questions truthfully. When questioned about their preferred mode of inquiry 82.5% said they would like to do similar surveys about sexual behaviour using a touch screen computer with ACASI.

Table 2 shows the results of the univariate models, developed to identify the groups of participants most likely to have a positive/negative user experience. Female participants (OR 1.68; 95% CI: 1.23–2.29) and those with a secondary (OR 2.00; 95% CI: 1.48–2.71) or tertiary education (OR 7.06; 95% CI: 1.62–30.82) were more likely to find the ACASI questionnaire easy to use, while those who were in the oldest age group (OR 0.37; 95% CI: 0.23–0.58) and the unemployed (OR 0.69; 95% CI: 0.47–1.01) were less likely to find it easy. Participants aged 25–39 (OR 2.10; 95% CI: 1.16–3.81) were more likely to view their experience as being confidential. Coloured participants were less likely to think our methods granted privacy (OR 0.46; 95% CI: 0.29–0.74). The oldest age group (OR 1.81; 95% CI: 1.06–3.09) and those with a secondary education (OR 1.53; 95% CI: 1.02–2.30) were the groups most likely to answer all questions truthfully.

**Table 1 Participant experience with ACASI mode of inquiry in cross-sectional sexual behaviour survey**

	n	%
<b>N</b>	798	
<b>Ease</b>		
Very easy	536	67.2
Somewhat easy	149	18.7
Somewhat difficult	84	10.5
Very difficult	28	3.5
I don't know	1	0.1
<b>Confidentiality</b>		
Very private	708	88.7
Somewhat private	61	7.6
No privacy	26	3.3
I don't know	3	0.4
<b>Truthfulness</b>		
Answered all truthfully	683	85.6
Answered most truthfully	84	10.5
Answered some truthfully, some dishonestly	13	1.6
Answered most dishonestly	9	1.1
Answered all dishonestly	7	0.9
I don't know	2	0.3
<b>Preferred mode of inquiry</b>		
Touch screen computer with ACASI	658	82.5
Researcher-administered verbal questionnaire	66	8.3
Self-administered written questionnaire	38	4.8
Telephonic survey	36	4.5
I don't know	0	0.0

Participants with a secondary education (OR 2.40; 95% CI: 1.64–3.50) were the most likely to choose touch screen ACASI survey as their preferred mode of inquiry, while those who belonged to the oldest age group (OR 0.44; 95% CI: 0.24–0.79) were less likely to choose this modality.

Table 3 provides a comparison of different reported sexual behaviours, by men and women, in our survey with three different FTFI-based surveys in southern Africa. In this comparison, we used only 503 of our original participants, creating a subset of participants with demographics most closely resembling the DHS sample population. For all sexual behaviours examined, our survey produces smaller male–female prevalence ratios than the DHSs in Lesotho, Swaziland, and Zimbabwe. The ratio of men to women reporting greater than 2 partners in the past year in our survey was 3.5 (95% CI: 2.3–5.2) compared to 10.6 (95% CI: 7.3–15.2) in Lesotho, 20.0 (95% CI: 7.1–72.9) in Swaziland, and 22.0 (95% CI: 9.4–49.0) in Zimbabwe. For the prevalence of greater than 2 lifetime partners, our survey produced a ratio of 1.2

**Table 2 Unadjusted odds ratios (ORs) for predictors of adverse interview experiences**

	Ease OR (95% CI)	Confidentiality OR (95% CI)	Truthfulness OR (95% CI)	Mode of inquiry OR (95% CI)
n	791	789	790	792
Age				
15–24 years	1.00	1.00	1.00	1.00
25–39 years	0.71 (0.45–1.14)	2.1 (1.16–3.81)	1.42 (0.85–2.37)	0.58 (0.32–1.07)
40–70 years	0.37 (0.23–0.58)	1.48 (0.84–2.60)	1.81 (1.06–3.09)	0.44 (0.24–0.79)
Gender				
Male	1.00	1.00	1.00	1.00
Female	1.68 (1.23–2.29)	1.40 (0.88–2.21)	1.02 (0.67–1.56)	1.03 (0.70–1.52)
Race				
Black	1.00	1.00	1.00	1.00
Coloured	1.03 (0.73–1.45)	0.46 (0.29–0.74)	0.88 (0.56–1.39)	1.34 (0.86–2.11)
Education				
None or primary	1.00	1.00	1.00	1.00
Secondary	2.00 (1.48–2.71)	1.39 (0.88–2.18)	1.53 (1.02–2.30)	2.40 (1.64–3.50)
Tertiary	7.06 (1.62–30.82)	0.89 (0.25–3.14)	0.89 (0.29–2.75)	2.00 (0.57–6.97)
Employment Status				
Employed	1.00	1.00	1.00	1.00
Unemployed	0.69 (0.47–1.01)	0.87 (0.50–1.52)	0.92 (0.56–1.51)	0.88 (0.55–1.39)

OR, Unadjusted Odds Ratio.  
 CI, Confidence Interval.

(95% CI: 1.0–1.3) compared to 2.1 (95% CI: 2.0–2.2) in Swaziland and 4.9 (95% CI: 4.6–5.2) in Zimbabwe. Our survey also produced a ratio of 2.6 (95% CI: 1.9–3.6) for respondents reporting a concurrent relationship in the past year (Swaziland 15.4, 95% CI: 10.7–21.7; Zimbabwe 11.3, 95% CI: 8.8–16.0). Finally, Table 3 also shows that the prevalence's of all reported behaviours are larger for men and women in our survey compared to the others.

### Discussion

We assessed the user-friendliness, privacy and truthfulness of an ACASI-based sexual behaviour survey that was administered in disadvantaged, urban communities

in Cape Town, South Africa. We found that an overwhelming majority of our participants preferred ACASI on touch screen computers to other modes of inquiry and that they answered all questions truthfully, probably owing in large part to participants' perceived ease of use and privacy.

Despite the predominately positive experience, our analysis indicates that more effort may be required to improve the user experience for certain subgroups of the South African population. For instance, respondents over the age of 40 were less likely to think ACASI was easy to use, which might explain why they were also less likely to prefer ACASI as a mode of inquiry. Indeed,

**Table 3 Percent of respondents reporting risk behaviours by gender in three DHS surveys using face-to-face interviews and the ACASI administered Cape Town Sexual Behaviour Survey**

Survey	Interview mode	Sexual behavior characteristics								
		> 2 Partners in past year			> 2 Lifetime partners			Had a concurrent partnership		
		Male (%)	Female (%)	PR	Male (%)	Female (%)	PR	Male (%)	Female (%)	PR
Cape Town Sexual Behavior Survey	ACASI	35.8	10.4	3.4	71.4	61.6	1.2	41.5	15.7	2.6
Lesotho DHS	FTFI	7.4	0.7	10.6	63.1	NA*	NA*	17.7	NA*	NA*
Swaziland DHS	FTFI	2.0	0.1	20.0	67.9	32.8	2.1	12.3	0.8	15.4
Zimbabwe DHS	FTFI	2.2	0.1	22.0	60.9	12.5	4.9	7.9	0.7	11.3

PR, male–female prevalence ratio.

ACASI, Audio Computer-Assisted Self-Interviewing.

FTFI, Face-to-Face Interviewing.

\* The DHS did not report prevalences for these female behaviors and therefore, the PR could not be calculated.

others have found that older generations are often intimidated by computers [28]. These results are not surprising given that computer literacy skills are only beginning to be taught in disadvantaged communities in South Africa. Moreover, unemployed participants found ACASI methods difficult to use probably due to the lack of computer exposure from a work environment. Since these groups of people found the survey challenging, it may have affected their recall of information due to the additional cognitive strain of navigating ACASI [13].

We also found that coloured participants were less likely to find our methods private. We suspect the reason for this is that a plurality of the coloured participants in our survey belonged to the oldest age group. McCallum and Peterson postulate that younger generations may find it more socially acceptable to share sensitive information due to changing views of privacy in the new technology age [13]. In fact, among Cape Town youth, there is a high use of mobile internet for instant messaging, digital media, and social networking [29], making their standards for privacy different from older generations. It should also be noted that in the multivariate analysis, adjusting for all of the potential predictors, the predictors did not change, only the size of the effects.

Our results also indicate that in this context, ACASI may produce more accurate results by partially removing some of the gendered SDB. In our survey the gender gap between reported sexual behaviours for men and women was narrowed, compared to the DHS that used FTFI techniques. Additionally, all of the sexual behaviours were reported more frequently in our survey than in the DHSs. This is particularly noticeable when looking at female reports of *> 2 partners in the past year and had a concurrent relationship*: less than 1% of women reported these behaviors in all DHS studies. This could be due to many different confounders or cultural differences between the two populations. However, we believe such a large difference in reported behaviours is most likely due to increased perception of privacy, which made our participants feel more at ease when answering intimate questions [30].

Our study had some limitations, which may affect the interpretation of our results. Firstly, we recognise that the questions we asked with respect to user experience, may also be subject to forms of bias. However, we believe that the bias would be minimal because typical forms of survey bias (e.g. inaccurate recall, social desirability, etc.) were probably not factors in answering these specific questions. Secondly, the positive results from the first two parts of our analysis may have been confounded by additional mechanisms we built into the survey to reduce bias, other than the interview mode. For one, we made use of a visual timeline to help respondents recall beginning and end dates of distinct episodes within each of their

relationships. These visual cues have been demonstrated to encourage internal consistency in reporting relationship histories [31,32]. Additionally, our survey was relatively short, asking only a few questions for each reported relationship episode and partner, thus minimizing fatigue bias compared to other long surveys, such as DHSs. While these are an overall strength of the survey, they may have contributed to the positive user experience, as much or more than the ACASI itself.

The greatest strength of this study is that it is one of few sexual behaviour surveys that measured the user experience of the instrument [33]. This allowed us not only to find a proxy for how accurate our results were, but it enabled us to quantitatively assess how appropriate our methods were for the communities where the survey was administered. By asking questions about user experience we were able to determine that old age groups and unemployed people in similar contexts may be at risk for having an adverse experience with ACASI and consequently inaccurately reporting their results. Going forward, this knowledge will allow researchers to target these at-risk groups and provide more assistance and training for them before conducting a survey with similar methods.

## Conclusions

The results of our study indicate that ACASI methods are not only suitable for disadvantaged, urban African contexts, but they may considerably reduce SDB in sexual behaviour surveys. By targeting groups at-risk for finding ACASI methods challenging, and providing them with additional support and training prior to taking the survey, the accuracy of data may be improved. However, future work needs to be done on how to best train these people to make them feel comfortable with computerized survey methods. Correspondingly, more work needs to be done on how computer literacy and numeracy training affect the validity of ACASI responses.

## Competing interests

The authors of this manuscript have no competing interests.

## Authors' contributions

RB, ND, MT, AW, NH and WD jointly designed the ACASI survey. RB coordinated data collection and wrote the first draft of the manuscript. FM conducted the statistical analysis. NH supervised the data analysis. All authors contributed during the editing process and approved the final, submitted manuscript.

## Acknowledgements

The authors wish to thank the survey respondents for their participation in this study. We are also indebted to the research assistants for countless hours of dedicated fieldwork and to the Desmond Tutu TB Centre at Stellenbosch University for their guidance in coordination and planning of this cross-sectional sexual behaviour survey. The Flemish Interuniversity Council (VLIR) and the Flemish Scientific Research Fund (FWO) provided financial support for the survey. WD, RB and AW received funding from the Canadian International Development Agency (CIDA). FM was funded by FWO and NH acknowledges support from the University of Antwerp scientific

chair in Evidence-Based Vaccinology, financed in 2009–2012 by an unrestricted donation from Pfizer.

#### Author details

<sup>1</sup>The South African Department of Science and Technology/National Research Foundation (DST/NRF) Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, 19 Jonkershoek Road, Stellenbosch 7600, South Africa. <sup>2</sup>International Centre for Reproductive Health, Ghent University, De Pintelaan 185, Ghent 9000, Belgium. <sup>3</sup>Center for Statistics, Interuniversity Institute of Biostatistics and Statistical Bioinformatics, Hasselt University, Diepenbeek Belgium. <sup>4</sup>Faculty of Medicine and Health Sciences, Ghent University, De Pintelaan 185, Ghent 9000, Belgium. <sup>5</sup>Centre for Health Economic Research and Modeling Infectious Diseases, Vaccine and Infectious Disease Institute, University of Antwerp, Wilrijk, Belgium.

Received: 15 October 2012 Accepted: 25 January 2013  
Published: 31 January 2013

#### References

- Schroder KE, Carey MP, Vanable PA: Methodological challenges in research on sexual risk behavior: I. Item content, scaling, and data analytical options. *Ann Behav Med* 2003, **26**(2):76–103.
- Crawford M, Popp D: Sexual double standards: a review and methodological critique of two decades of research. *J Sex Res* 2003, **40**(1):13–26.
- Bolly MC, Alary M, Baggaley RF: Neglected issues and hypotheses regarding the impact of sexual concurrency on HIV and sexually transmitted infections. *AIDS Behav* 2012, **16**(2):304–311.
- Wellings K, Collumbien M, Slaymaker E, Singh S, Hodges Z, Patel D, Bajos N: Sexual behaviour in context: a global perspective. *Lancet* 2006, **368**(9548):1706–1728.
- Zaba B, Pisani E, Slaymaker E, Ties Boerma J: Age at first sex: understanding recent trends in African demographic surveys. *Sex Transm Infect* 2004, **80**(Suppl 1):i128–i135.
- Clark S, Kabiru C, Zulu E: Do Men and women report their sexual partnerships differently? evidence from Kisumu, Kenya. *Int Perspect Sex R H* 2011, **37**(4):181–190.
- Mnyika KS, Klepp KI, Kvale G, OleKingori N: Determinants of high-risk sexual behaviour and condom use among adults in the Arusha region, Tanzania. *Int J STD AIDS* 1997, **8**(3):176–183.
- Nnko S, Boerma JT, Urassa M, Mwaluko G, Zaba B: Secretive females or swaggering males? An assessment of the quality of sexual partnership reporting in rural Tanzania. *Soc Sci Med* 2004, **59**(2):299–310.
- Prata N, Vahidnia F, Fraser A: Gender and relationship differences in condom use among 15-24-year-olds in Angola. *Int Fam Plan Perspec* 2005, **31**(4):192–199.
- Buve A, Lagarde E, Carael M, Rutenberg BJRF, Glynn JR, Laourou MEA, Chege J, Sukwa T, et al: Interpreting sexual behaviour data: validity issues in the multicentre study on factors determining the differential spread of HIV in four African cities. *AIDS* 2001, **15**(suppl 4):S117–S126.
- Meekers D, Klein M: Determinants of condom use among young people in urban Cameroon. *Stud Family Plann* 2002, **33**(4):335–346.
- Cordero-Coma J, Breen R: HIV prevention and social desirability: husband-wife discrepancies in reports of condom Use. *J Marriage Fam* 2013, **74**:601–613.
- McCallum EB, Peterson ZD: Investigating the impact of inquiry mode on self-reported sexual behavior: theoretical considerations and review of the literature. *J Sex Res* 2012, **49**(2–3):212–226.
- Jaya, Hindin MJ, Ahmed S: Differences in young people's reports of sexual behaviors according to interview methodology: a randomized trial in India. *Am J Public Health* 2008, **98**(1):169–174.
- Rathod SD, Minnis AM, Subbliah K, Krishnan S: ACASI and face-to-face interviews yield inconsistent estimates of domestic violence among women in India: The Samata Health Study 2005–2009. *J Interpers Violence* 2011, **26**(12):2437–2456.
- Rose E, Diclemente RJ, Wingood GM, Sales JM, Latham TP, Crosby RA, Zenilman J, Melendez J, Hardin J: The validity of teens' and young adults' self-reported condom use. *Arch Pediatr Adolesc Med* 2009, **163**(1):61–64.
- Dolezal C, Marhefka SL, Santamaria EK, Leu CS, Brackis-Cott E, Mellins CA: A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Arch Sex Behav* 2012, **41**(2):401–410.
- Phillips AE, Gomez GB, Bolly MC, Garnett GP: A systematic review and meta-analysis of quantitative interviewing tools to investigate self-reported HIV and STI associated behaviours in low- and middle-income countries. *Int J Epidemiol* 2010, **39**(6):1541–1555.
- Kissinger P, Rice J, Farley T, Trim S, Jewitt K, Margavio V, Martin DH: Application of computer-assisted interviews to sexual behavior research. *Am J Epidemiol* 1999, **149**(10):950–954.
- Hewett P, Mensch B, Enulkar A: Consistency in the reporting of sexual behavior by adolescent girls in Kenya: a comparison of interviewing methods. *Sex Transm Infect* 2004, **80**(Suppl 2):i43–i48.
- Central Statistical Office (CSO) S, Inc M: *Swaziland Demographic and Health Survey 2006–07*. Demographic and Health Survey, Mbabane, Swaziland: Central Statistical Office Macro International Inc; 2008.
- Central Statistical Office (CSO) Z, Inc M: *Zimbabwe Demographic and Health Survey 2005–06*. Calverton, Maryland: CSO and Macro International Inc; 2007.
- Ministry of Health and Social Welfare (MOHSW) L, (BOS) Bos, Macro O: *Lesotho Demographic and Health Survey 2004*. Calverton, Maryland: MOH, BOS, and ORC Macro; 2005.
- Delva W, Beauclair R, Welte A, Vansteelandt S, Hens N, Aerts M, du Toit E, Beyers N, Temmerman M: Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol. *BMC Public Health* 2011, **11**:616.
- The American Association for Public Opinion Research: *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. AAPOR, 7th edition; 2011.
- Mishra V, Assche B-V: *Concurrent Sexual Partnerships and HIV Infection: Evidence from National Population-Based Surveys*. DHS Working Papers No 62. Calverton, Maryland: Macro International Inc; 2009.
- R Development Core Team: *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing; 2005.
- Feigelson ME, Dwight SA: Can asking questions by computer improve the candor of responding? A meta-analytic perspective. *Consulting Psychol J: Pract Res* 2000, **52**(4):248–255.
- Kreutzer T: *Generation Mobile: Online and Digital Media Usage on Mobile Phones among Low-Income Urban Youth in South Africa*. Cape Town: University of Cape Town; 2009.
- Lanhaug LF, Sherr L, Cowan FM: How to improve the validity of sexual behavior reporting: systematic review of questionnaire delivery modes in developing countries. *Trop Med Int Health* 2010, **15**(3):362–381.
- Kabiru CW, Luke N, Izugbara CO, Zulu E: The Correlates of HIV Testing and Impacts on Sexual Behaviour: evidence from a life history study of young people in Kisumu, Kenya. *BMC Public Health* 2010, **10**:412.
- Luke N, Clark S, Zulu E: *The Relationship History Calendar: Improving Sexual Behaviour Data among Youth in Developing Country Settings*. New Orleans, USA: Population Association of America (PAA) Conference; 2008.
- Estes L, Lloyd L, Teti M, Raja S, Bowleg L: Perceptions of Audio Computer-Assisted Self-Interviewing (ACASI) among Women in an HIV-Positive Prevention Program. *PLoS One* 2010, **5**(2):e9149.

doi:10.1186/1471-2288-13-11

Cite this article as: Beauclair et al.: Evaluating audio computer assisted self-interviews in urban south African communities: evidence for good suitability and reduced social desirability bias of a cross-sectional survey on sexual behaviour. *BMC Medical Research Methodology* 2013 **13**:11.



## Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol

Delva *et al.*

STUDY PROTOCOL

Open Access

# Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol

Wim Delva<sup>1,2†</sup>, Roxanne Beauclair<sup>1\*†</sup>, Alex Welte<sup>1</sup>, Stijn Vansteelandt<sup>3</sup>, Niel Hens<sup>4,5</sup>, Marc Aerts<sup>4</sup>, Elizabeth du Toit<sup>6</sup>, Nulda Beyers<sup>6</sup> and Marleen Temmerman<sup>2</sup>

## Abstract

**Background:** Crucial connections between sexual network structure and the distribution of HIV remain inadequately understood, especially in regard to the role of concurrency and age disparity in relationships, and how these network characteristics correlate with each other and other risk factors. Social desirability bias and inaccurate recall are obstacles to obtaining valid, detailed information about sexual behaviour and relationship histories. Therefore, this study aims to use novel research methods in order to determine whether HIV status is associated with age-disparity and sexual connectedness as well as establish the primary behavioural and socio-demographic predictors of the egocentric and community sexual network structures.

**Method/Design:** We will conduct a cross-sectional survey that uses a questionnaire exploring one-year sexual histories, with a focus on timing and age disparity of relationships, as well as other risk factors such as unprotected intercourse and the use of alcohol and recreational drugs. The questionnaire will be administered in a safe and confidential mobile interview space, using audio computer-assisted self-interview (ACASI) technology on touch screen computers. The ACASI features a choice of languages and visual feedback of temporal information. The survey will be administered in three peri-urban disadvantaged communities in the greater Cape Town area with a high burden of HIV. The study communities participated in a previous TB/HIV study, from which HIV test results will be anonymously linked to the survey dataset. Statistical analyses of the data will include descriptive statistics, linear mixed-effects models for the inter- and intra-subject variability in the age difference between sexual partners, survival analysis for correlated event times to model concurrency patterns, and logistic regression for association of HIV status with age disparity and sexual connectedness.

**Discussion:** This study design is intended to facilitate more accurate recall of sensitive sexual history data and has the potential to provide substantial insights into the relationship between key sexual network attributes and additional risk factors for HIV infection. This will help to inform the design of context-specific HIV prevention programmes.

## Background

In South Africa, 16.9% of the 15 to 49 years old population is infected with HIV [1]. The high prevalence of HIV in South Africa can be at least partially attributed to widespread poverty and labour migration from rural areas to urban centres that has subsequently upset family structures and disrupted psychosocial support

systems [2-4]. Furthermore, continued segregation of communities along racial lines is a remnant of the old Apartheid political system and it has been suggested that this has mostly confined the epidemic to black and to a lesser extent coloured communities (the term used to describe the racially mixed descendants of Europeans, indigenous populations and slaves from South and East Asia) [5]. While the overall HIV prevalence is similar among men and women, more than 1 in 5 women (21.1%) between 20 and 24 years old is HIV-infected compared to only 1 in 20 (5.1%) men in the same age group [6]. Besides physiological factors that may

\* Correspondence: rbeauclair@sun.ac.za

† Contributed equally

<sup>2</sup>South African Centre for Epidemiological Modelling & Analysis, Stellenbosch University, 19 Jonkershoek Road, Stellenbosch 7600, South Africa  
Full list of author information is available at the end of the article

increase young women's susceptibility to HIV acquisition, their frequent engagement in age-disparate relationships is assumed to impose additional risk of HIV infection. Indeed, recent data suggests that women in sexual relationships with men five or more years older are more likely to be HIV infected [1,7,8].

For young women in South Africa, an older partner is more likely to expose them to the virus, since the HIV prevalence in men increases with age, peaking in their early thirties [6]. Moreover, older men who have sexual relationships with younger women facilitate HIV transmission because they are less likely to use condoms consistently than men from the same age cohorts who do not have age disparate relationships [9,10]. Lastly, older male partners of young women frequently partake in multiple, concurrent relationships with casual partners while maintaining 'long-term' partnerships [11].

In concert with age-disparate relationships, concurrent relationships - overlapping sexual partnerships where sexual intercourse with one partner occurs between 2 acts of intercourse with another partner - have been purported by several authors to facilitate the spread of HIV [12-18]. Contrary to this assertion, are other studies that question the evidence for concurrency as a key driver of the HIV epidemics in Southern Africa [19-21]. The protective effect that polygamy has against acquiring HIV serves as counterevidence for a positive concurrency-HIV correlation [22]. The level of connectedness of individuals to the broader sexual network, rather than the number of overlapping relationships, may be a more important determinant of the spread of HIV in communities. Having concurrent relationships, changing partners frequently, and choosing partners from a wide range of age groups and geographical locations all increase one's sexual connectedness. The degree to which these behaviours are correlated and the strength of association between these behaviours and one's HIV status remains incompletely understood.

To this end, our study will be located in three disadvantaged communities in Cape Town, South Africa. The primary objective is to determine whether HIV status is associated with age-disparity and sexual connectedness. The secondary objectives of the study are to ascertain behavioural and socio-demographic predictors of the egocentric and community sexual network structures. To accomplish these objectives we will implement novel methods that will address many of the typical challenges of obtaining unbiased sexual behaviour information.

## Methods/Design

### Study Design

We will undertake a cross-sectional, sexual behaviour survey. This study is registered with and approved by the Stellenbosch University Health Research Ethics

Committee (N11/03/093) and it is a subsidiary study to the Zambia South Africa TB and AIDS Reduction study (ZAMSTAR) (ISRCTN36729271), a community randomised trial aimed at reducing the prevalence of tuberculosis in communities with a high burden of TB and HIV, by novel public health interventions.

### Study Setting

The study outlined here will take place in three peri-urban, disadvantaged communities in the greater Cape Town area with a high burden of HIV. These communities represent one predominantly black community and two racially diverse communities. They were chosen in order to gather a representative sample of the different races most affected by HIV in South Africa.

### Participants

The final primary outcome measurements of the ZAMSTAR study were collected through an HIV and TB prevalence survey, conducted in 2010. In this prevalence survey, participants provided written informed consent to complete a questionnaire on basic demographic information and medical history, to provide sputum for TB culture, and to have repeated visits by the research team. Through an additional consent procedure, participants gave consent for HIV counselling and testing. Our study will randomly select 500 previous ZAMSTAR participants in each of the three communities. In the ZAMSTAR survey, some participants chose not to have HIV testing done. For this reason, 20% of our sample per study site (100/500) will consist of participants who did not consent for HIV testing. By sampling both participants who did, and who did not, consent to an HIV test, we may be able to detect selection bias.

Study participants will be included if they are between the ages of 18 and 64 at the time of the ZAMSTAR survey, able to give informed consent, and agree to complete the study questionnaire. Individuals who do not speak English, Afrikaans or isiXhosa, or who are enrolled in any other ongoing study, will be excluded.

### Measurement Instruments

A touch screen questionnaire, utilizing an audio computer-assisted self-interviewing (ACASI) application has been developed to facilitate the collection of sexual history data. The respondents will wear headphones to hear questions in their choice of English, Afrikaans, or isiXhosa, and simultaneously be able to read the question and select their answers on a 22-inch touch screen monitor. This question format was derived from a UNAIDS "Best Practice" questionnaire of sexual partnership data [23], as well as the Relationship History Calendar (RHC) and an Events History Calendar (EHC) [24]. These calendars have been previously validated and

tested in adolescent and adult respondents in sub-Saharan Africa [24,25] and the United States [26]. The modified RHC used in this study will collect detailed retrospective data on sexual histories of the participants for the year preceding the survey, for a maximum of 5 'main' sexual partners and 15 'casual' sexual partners. Before starting the actual survey, participants will view a short demonstration video and will practise using ACASI and the touch screen by answering a series of example questions.

The questionnaire is set up to ask questions along a temporal trajectory. It will begin by asking the participant basic demographic information and then proceed to ask if the participant had a main sexual partner one year ago and if that relationship is still ongoing. The participant will be able to indicate, on a touch screen timeline, the periods they were in this relationship. For each relationship, we ask whether or not the participant or his/her partner used drugs or alcohol at their first intercourse. To gauge the level of spatial connectedness, we ask how long the participant usually travelled to their partner and his/her degree of proximity. For each distinct period that the respondent engaged in sexual activity with a particular partner, measured in weeks, additional questions will be asked about the frequency of intercourse and condom use. This series of questions will be asked for each main sexual partner, as well as the casual sexual partners. The onset, dissolution, and duration of each relationship will be displayed on the touch screen timeline for the participant to see, using different colours for each partner. At the end of the questionnaire, participants will be asked how many partners they have ever had. In addition to the sexual behaviour questionnaire we added three questions to assess the ease of use, perceived confidentiality and self-reported truthfulness when answering questions in this survey. Finally, we ask them to indicate their preferred medium of answering questions about sexual behaviour: touch screen computer with ACASI, researcher-administered verbal questionnaire, self-administered written questionnaire, or telephonic survey.

All questions were developed and refined after conducting cognitive interviews with six people who were representatives of the study communities. These interviews assessed the clarity, comprehensibility, and cultural sensitivity of the proposed questions, allowing us to improve the phrasing of questions, incorporate meaningful slang, and define suitable categorical answer options that maximize the precision and accuracy of responses. Furthermore, the cognitive interviews allowed us to gauge the anticipated community response to a survey utilizing ACASI. Our interlocutors responded positively to the idea of answering questions about their sexual histories in a completely anonymous way and

articulated that the communities would embrace a survey that could accomplish this.

#### **Survey administering and data management**

The senior data manager will have access to ZAMSTAR study participant information and will therefore be able to obtain names and addresses for each individual. Consent forms will be printed that include a new unique barcode - different from but linked to the barcode allocated in the ZAMSTAR study - along with names and addresses of the individuals. The barcodes from the ZAMSTAR study will not be printed on these consent forms. Therefore, none of the study staff, except the data manager, will be able to link the HIV test results obtained in the ZAMSTAR survey to our survey data.

Residents of the community, who were confirmed with names, addresses and dates of birth to be previous ZAMSTAR participants, will be asked if they consent to participation after reading an information sheet and listening to verbal information given by the research assistants in their home language. The research assistants will make it known that the anonymised HIV test results from the ZAMSTAR study will be accessed and used in the analysis. If the participant consents to the survey and signs the consent form, he/she will be escorted to a camper van - located in a safe part of the neighbourhood - where the survey will commence. The camper van has been refurbished to provide a private and safe office space, containing a partition, in which two participants can take the survey simultaneously at their own desk, chair and computer. Inside the van, the research assistant will scan the barcode on the consent form into the touch screen computer.

All subsequent answers to the touch screen questionnaire will be answered in private by the participant only. No one will be able to see their responses, including the research assistants who will be waiting outside of the van or in the driver's cab. They will remain close by at all times in the event that a participant requests their assistance. While two participants take the survey, two additional research assistants will be out recruiting the next two participants from the same community. The participant's name will not be asked in the questionnaire, hence the senior data manager will be the only person who can link the name from the consent form via the unique barcode. The touch screen questionnaire data will be uploaded from the laptop onto the central database designed for this study. All data will be backed up onto a second server. The consent forms will be stored in locked cabinets at the study centre.

To ensure the quality of the data collection process, the fieldwork coordinator will randomly select 5% of the completed consent forms every week and do a home visit to verify that the participant was in fact enrolled in

the study and to confirm the signature on the consent form.

#### HIV Status

The HIV status of participants was determined in the 2010 ZAMSTAR Tuberculosis Prevalence Survey. In that study, for those who consented to HIV testing, Abbott Determine HIV-1/2 screening tests were used, and a second, confirmatory test was conducted for those who tested positive. The data manager of this study is the only person to have access to the ZAMSTAR HIV test results and after the completion of our survey, he will link these test results to the new survey data.

#### Statistical Analyses

In an initial descriptive analysis, the prevalence of HIV infection and the point prevalence of concurrent relationships and age-disparate relationships will be calculated, along with the average number of partners per year, the average frequency of sexual intercourse, the frequency of condom use for each of the reported relationships, condom use at last sexual intercourse with each of the reported partners, the average age difference between individuals and their partners and the variance of these age differences in the study population. Since many individuals are likely to report on more than one relationship, linear mixed-effects models will be used to analyze the inter- and intra-subject variability in the age difference between sexual partners, and covariates associated with large age differences. To take into account the fact that multiple sexual partnerships may co-exist, and the timing of partnership initiation and dissolution are correlated, or dependent on each other, survival analysis for correlated event times will be conducted to model concurrency patterns, partnership durations and rates of partnership initiation and termination. This type of survival analysis also accommodates right-censored observations where current partnership durations are only known up to the time of the survey. Marginal survival models for correlated event times will be used to enable comparisons between genders or other covariate groupings.

Next, logistic regression models for clustered data (due to respondents reporting on the characteristics of multiple partners, constituting repeated measures) will be fit to the data to determine whether HIV status is associated with age-disparity and sexual connectedness. Age-disparity and sexual connectedness will be operationalised using the following individual and community characteristics: (a) the mean difference between the age of an individual and the age of his/her partners; (b) the variability in age difference between an individual and his/her partners; (c) being engaged in a concurrent relationship; (d) having ever engaged in concurrent

relationships; (e) the number of past and present concurrent relationships; (f) the cumulative overlapping time engaged in concurrent relationships; (g) the number of lifetime sexual partners; (h) spatial proximity to sexual partners; (i) the population mean age differences between individuals and their partners; (j) the population variability in age difference between individuals and their partners; (k) point prevalence of concurrency; (l) prevalence of having ever engaged in concurrent relationships; (m) population mean number of past and present concurrent relationships; (n) the population per capita cumulative overlapping time engaged in concurrent relationships; (o) population mean number of lifetime partners; (p) population spatial assortativeness (i.e. choosing partners from one's own community).

In a supplementary analysis, socio-demographic and behavioural predictors of egocentric and community sexual network structures will be ascertained. Potential risk factors under consideration are: race, gender, proximity to city centre, education level, socio-economic status, age, religion, and alcohol and drug use at first intercourse. Confirmed risk factors will be added to the models as confounders in the primary analyses and adjusted associations for age-disparity, concurrency and HIV status at the individual and cluster level will be calculated.

#### Sample size calculations

For an alpha level of 5%, a design effect of 2.0 (effect of clustered data rather than independent random sample) and a prevalence of HIV infection of 15%, with the width of the 95% confidence interval at 6% (i.e. +/-3%), 1089 study participants with HIV test results are needed. Building a 10% margin for inconsistent and incomplete data, we will aim to administer the survey to 1200 study participants for whom HIV test results are available. Additionally, we will include 300 study participants (20% of the total sample) who did not opt for HIV testing, to investigate whether non-universal consent for HIV testing may have introduced selection bias.

#### Discussion

This study will describe age-disparity and sexual connectedness in disadvantaged communities around Cape Town, South Africa. These aspects of sexual networks appear to be important for the spread of HIV, yet their complex patterns and the way in which they are driving HIV transmission are insufficiently understood. One of the fundamental challenges to understanding the role of sexual network structure in the spread of HIV is obtaining valid sexual history data. We intend to address this challenge by using a survey design that attempts to minimize social desirability bias and maximize confidentiality.

Specifically, we created a mobile office space where participants are free to answer questions away from their homes, where they often encounter many familial disruptions and potential eavesdroppers that could influence how they answer questions about their sexual behaviours. Secondly, we decided to use ACASI methods, which have successfully been used in previous sexual behaviour surveys to reduce social desirability bias. This bias is introduced when participants are intimidated by the researchers or fieldwork assistants and experience pressure to provide socially accepted answers, not necessarily in line with the truth. Studies comparing ACASI methods with face-to-face interviews, suggest that participants are more likely to report high-risk sexual behaviours while using the ACASI [27-31].

In addition to this, our study will also attempt to curb recall bias about certain sexual partners and behaviours. Events History Calendar (EHC) methods have the potential to unlock autobiographical memory better than traditional survey methods [32] because they facilitate the recollection of complex life history data [33] and they supply a constant visual cue that enhances a participant's ability to precisely recall timing of events [32,34].

This survey will expand the body of knowledge on sexual and behavioural determinants of HIV infection, with special focus on the role of age disparity and sexual connectedness. Since the population impact of an HIV prevention method does not only depend on its intrinsic efficacy, but also on the structure of the sexual network in which it is introduced, our study hopes to contribute to the development of appropriate HIV preventative strategies that take the local sexual network structure into account.

#### Acknowledgements

This study is funded by the Flemish Scientific Research Fund (FWO) and the Flemish Interuniversity Council (VLIR).

#### Author details

<sup>1</sup>South African Centre for Epidemiological Modelling & Analysis, Stellenbosch University, 19 Jonkershoek Road, Stellenbosch 7600, South Africa.  
<sup>2</sup>International Centre for Reproductive Health, Ghent University, De Pintelaan 185, 9000 Gent, Belgium.  
<sup>3</sup>Department of Applied Mathematics and Computer Science, Ghent University, Krijgslaan 281, S9, 9000 Gent, Belgium.  
<sup>4</sup>Interuniversity Institute for Biostatistics and statistical Bioinformatics (I-BIOSTAT), Hasselt University, Agoralaan - building D, 3590 Diepenbeek, Belgium.  
<sup>5</sup>Centre for Health Economics Research and Modeling Infectious Diseases (CHERMID), Centre for the Evaluation of Vaccination, Vaccine & Infectious Disease Institute, University of Antwerp, Campus Drie Eiken CDE, Universiteitsplein 1, 2610 Antwerp, Belgium.  
<sup>6</sup>Desmond Tutu TB Centre, Department of Paediatrics and Child Health, Faculty of Health Sciences, Stellenbosch University, Francie Van Zyl Road, Cape Town 7507, South Africa.

#### Authors' contributions

WD, AW, ED, NB and MT were responsible for the conceptual design of the study. All authors participated in revisions to the study design and approved the final study design. RB and WD were involved in drafting of the manuscript, all other authors were involved in overall revision of the

manuscript. All authors are involved in the implementation of the project, and have read and approved the final manuscript.

#### Competing interests

The authors declare that they have no competing interests.

Received: 14 July 2011 Accepted: 2 August 2011

Published: 2 August 2011

#### References

1. Shisana O, Rehle T, Simbayi LC, Parker W, Zuma K, Bhana A, Connolly C, Jooste S, Pillay V: South African national HIV prevalence, HIV incidence, behaviour and communication survey, 2005. Cape Town: HSRC Press; 2005.
2. Coffee M, Garnet G, Millo M, Voeten H, Chandilwana S, Gregson S: Patterns of movement and risk of HIV infection in rural Zimbabwe. *Journal of Infectious Diseases* 2005, 191(Suppl 1):S159-167.
3. Hargrove J: Migration, mines and mores: the HIV epidemic in southern Africa. *S Afr J Sci* 2008, 104:53-61.
4. Rees D, Murray J, Nelson G, Sonnenberg P: Oscillating migration and the epidemics of silicosis, tuberculosis, and HIV infection in South African gold miners. *Am J Ind Med* 2009.
5. Kenyon C, Dlamini S, Boulie A, White R, Badr M: A network-level explanation for the differences in HIV prevalence in South Africa's racial groups. *African Journal of AIDS Research* 2009, 8(3):243-254.
6. Shisana O, Rehle T, Simbayi L, Zuma K, Jooste S, Pillay-van-Wyk V, Mbelle N, Van Zyl J, Parker W, Zungu N, et al: South African national HIV prevalence, incidence, behaviour and communication survey 2008: A turning tide among teenagers? Cape Town: Human Sciences Research Council; 2009.
7. Kelly RJ, Gray RH, Sewankambo NK, Serwadda D, Wabwire-Mangen F, Lutalo T, Wawer MJ: Age differences in sexual partners and risk of HIV-1 infection in rural Uganda. *J Acquir Immune Defic Syndr* 2003, 32(4):446-451.
8. Sa Z, Larsen U: Gender inequality increases women's risk of hiv infection in Moshi, Tanzania. *J Biosoc Sci* 2008, 40(4):505-525.
9. Bankole A, Ahmed FH, Neema S, Quedraogo C, Konyani S: Knowledge of correct condom use and consistency of use among adolescents in four countries in Sub-Saharan Africa. *Afr J Reprod Health* 2007, 11(3):197-220.
10. Luke N: Confronting the 'sugar daddy' stereotype: age and economic asymmetries and risky sexual behavior in urban Kenya. *Int Fam Plan Perspect* 2005, 31(1):6-14.
11. Chopra M, Townsend L, Johnston L, Mathews C, Tomlinson M, O'Bra H, Kendall C: Estimating HIV prevalence and risk behaviors among high-risk heterosexual men with multiple sex partners: use of respondent-driven sampling. *J Acquir Immune Defic Syndr* 2009, 51(1):72-77.
12. Halperin DT, Epstein H: Concurrent sexual partnerships help to explain Africa's high HIV prevalence: implications for prevention. *Lancet* 2004, 364(9428):4-6.
13. Hudson CP: Concurrent partnerships could cause AIDS epidemics. *Int J STD AIDS* 1993, 4(5):249-253.
14. Kretzschmar M, Morris M: Measures of concurrency in networks and the spread of infectious disease. *Math Biosci* 1996, 133(2):165-195.
15. Mah TL, Halperin DT: Concurrent sexual partnerships and the HIV epidemics in Africa: evidence to move forward. *AIDS Behav* 14(1):11-16, discussion 34-17.
16. Morris M, Kretzschmar M: Concurrent partnerships and the spread of HIV. *AIDS* 1997, 11(5):641-648.
17. Rothenberg R: HIV transmission networks. *Curr Opin HIV AIDS* 2009, 4(4):260-265.
18. Watts C, May R: The influence of concurrent partnerships on the dynamics of HIV/AIDS. *Math Biosci* 1992, 108(1):89-104.
19. Lagarde E, Auvert B, Carael M, Laourou M, Ferry B, Akam E, Sukwa T, Morison L, Maury B, Chege J, et al: Concurrent sexual partnerships and HIV prevalence in five urban communities of sub-Saharan Africa. *AIDS* 2001, 15(7):877-884.
20. Mishra V, Binami-Van Assche S: Concurrent sexual partnerships and HIV infection: evidence from national population-based surveys. *DHS Working papers* Calverton: Macro International Inc; 2009.
21. Lurie MN, Rosenthal S: Concurrent Partnerships as a Driver of the HIV Epidemic in Sub-Saharan Africa? The Evidence is Limited. *AIDS Behav* 2009.

22. Reniers G, Watkins S: Polygyny and the spread of HIV in sub-Saharan Africa: a case of benign concurrency. *AIDS* 2009.
23. UNAIDS: Looking Deeper into the HIV Epidemic: A Questionnaire for Tracing Sexual Networks. Edited by: HIV/AIDS JUNPo 1998.
24. Luke N, Clark S, Zulu E: The Relationship History Calendar: Improving Sexual Behaviour Data among Youth in Developing Country Settings. *Population Association of America (PAA) Conference* New Orleans, USA; 2008.
25. Kabiru CW, Luke N, Izugbara CO, Zulu E: The Correlates of HIV Testing and Impacts on Sexual Behaviour: evidence from a life history study of young people in Kisumu, Kenya. *BMC Public Health* 2010, **10**.
26. Martyn KK, Martin R: Adolescent Sexual Risk Assessment. *Journal of Midwifery & Women's Health* 2003, **48**(3):213-219.
27. Harris KM, Cheng MM: Dating Major Life Events using an CAPI/CASI Event History Calendar in Add Health. *Conference on the Event History Calendar Method* Washington DC; 2007.
28. Hewett PC, Mensch BS, Erukhar AS: Consistency in the reporting of sexual behaviour by adolescent girls in Kenya: a comparison of interviewing methods. *Sex Transm Infect* 2004, **80**(Suppl 2):ii43-48.
29. Turner AN, De Kock AE, Meehan-Ritter A, Blanchard K, Sebola MH, Hoosen AA, Coetzee N, Ellertson C: Many vaginal microbicide trial participants acknowledged they had misreported sensitive sexual behavior in face-to-face interviews. *J Clin Epidemiol* 2009, **62**(7):759-765.
30. van der Elst EM, Okuku HS, Nakamya P, Muhaari A, Davies A, McClelland RS, Price MA, Smith AD, Graham SM, Sanders EJ: Is audio computer-assisted self-interview (ACASI) useful in risk behaviour assessment of female and male sex workers, Mombasa, Kenya? *PLoS One* 2009, **4**(5):e5340.
31. Mensch BS, Hewett PC, Abbott S, Rankin J, Littlefield S, Ahmed K, Cassim N, Patel S, Ramjee G, Palanee T, et al: Assessing the reporting of adherence and sexual activity in a simulated microbicide trial in South Africa: an interview mode experiment using a placebo gel. *AIDS Behav* 2011, **15**(2):407-421.
32. Belli RF: The structure of autobiographical memory and the event history calendar: potential improvements in the quality of retrospective reports in surveys. *Memory* 1998, **6**(4):383-406.
33. Freedman D, Thornton A, Camburn D, Alwin D, Young-DeMarco L, (Eds.): *In The life history calendar: A technique for collecting retrospective data. Volume 18.* San Francisco: Jossey-Bass; 1988.
34. Caspi A, Moffitt TE, Thornton A, Freedman D, Amell JW, Harrington H, Smijers J, Silva PA: The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research* 1995, **6**:101-114.

#### Pre-publication history

The pre-publication history for this paper can be accessed here:  
<http://www.biomedcentral.com/1471-2458/11/616/prepub>

doi:10.1186/1471-2458-11-616

Cite this article as: Delva et al.: Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol. *BMC Public Health* 2011 **11**:616.

#### Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)



# APPENDIX C

Delva W et al. *Journal of the International AIDS Society* 2013, 16:18034  
<http://www.jiasociety.org/Index.php/jias/article/view/18034> | <http://dx.doi.org/10.7448/IAS.16.1.18034>



## Research article

### Coital frequency and condom use in monogamous and concurrent sexual relationships in Cape Town, South Africa

Wim Delva<sup>6,1,2</sup>, Fei Meng<sup>2,3</sup>, Roxanne Beauclair<sup>1</sup>, Nele Deprez<sup>4</sup>, Marleen Temmerman<sup>2,4</sup>, Alex Witte<sup>1</sup> and Niel Hens<sup>3,5</sup>

<sup>6</sup>Corresponding author: Wim Delva, SACEMA, c/o STIAS, Private Bag X1, Matieland 7602, Stellenbosch, Republic of South Africa. Tel: +27218082589. Fax: +27218082586. (Wim.Delva@ugent.be)

#### Abstract

**Introduction:** A decreased frequency of unprotected sex during episodes of concurrent relationships may dramatically reduce the role of concurrency in accelerating the spread of HIV. Such a decrease could be the result of coital dilution – the reduction in per-partner coital frequency from additional partners – and/or increased condom use during concurrency. To study the effect of concurrency on the frequency of unprotected sex, we examined sexual behaviour data from three communities with high HIV prevalence around Cape Town, South Africa.

**Methods:** We conducted a cross-sectional survey from June 2011 to February 2012 using audio computer-assisted self-interviewing to reconstruct one-year sexual histories, with a focus on coital frequency and condom use. Participants were randomly sampled from a previous TB and HIV prevalence survey. Mixed effects logistic and Poisson regression models were fitted to data from 527 sexually active adults reporting on 1210 relationship episodes to evaluate the effect of concurrency status on consistent condom use and coital frequency.

**Results:** The median of the per-partner weekly average coital frequency was 2 (IQR: 1–3), and consistent condom use was reported for 36% of the relationship episodes. Neither per-partner coital frequency nor consistent condom use changed significantly during episodes of concurrency (aIRR = 1.05; 95% confidence interval (CI): 0.99–1.24 and aOR = 1.01; 95% CI: 0.38–2.68, respectively). Being male, coloured, having a tertiary education, and having a relationship between 2 weeks and 9 months were associated with higher coital frequencies. Being coloured, and having a relationship lasting for more than 9 months, was associated with inconsistent condom use.

**Conclusions:** We found no evidence for coital dilution or for increased condom use during concurrent relationship episodes in three communities around Cape Town with high HIV prevalence. Given the low levels of self-reported consistent condom use, our findings suggest that if the frequency of unprotected sex with each of the sexual partners is sustained during concurrent relationships, HIV-positive individuals with concurrent partners may disproportionately contribute to onward HIV transmission.

**Keywords:** coital dilution; condom use; concurrency; HIV; South Africa; sexual behaviour; sex frequency.

Received 28 August 2012; Revised 1 February 2013; Accepted 27 February 2013; Published 24 April 2013

Copyright: © 2013 Delva W et al; licensee International AIDS Society. This is an open access article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) Licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

#### Introduction

Concurrent relationships have been defined by the Working Group on Measuring Concurrent Sexual Partnerships of the UNAIDS Reference Group on Estimates, Modelling, and Projections as “overlapping sexual partnerships in which sexual intercourse with one partner occurs between two acts of intercourse with another partner” [1]. The importance of concurrency in driving HIV transmission in hyperendemic settings remains controversial. While some have argued, primarily using modelling studies, that concurrency is a strong facilitator of HIV transmission, or even an essential driver for sustained HIV epidemics [2–4], others have dismissed the concurrency hypothesis, because of perceived flaws in the structure and assumptions of the models used [5] and missing empirical evidence for causal links between levels of concurrency and the local or national HIV prevalence [6–10].

Recently, Sawers *et al.* concluded that the role of concurrency in accelerating the spread of HIV is dramatically reduced by coital dilution – the reduction in per-partner coital frequency that accompanies the acquisition of additional partners [11]. In general, a decreased frequency of unprotected sex during episodes of concurrent relationships would reduce the transmission-facilitating effect of concurrency. Such a decrease could be the result of coital dilution and/or increased condom use during concurrency [12,13].

Despite the large number of sexual behaviour surveys that have investigated condom use, sex frequency and concurrency in settings with high HIV prevalence, few analyses have specifically focused on condom use and sex frequency in concurrent versus monogamous relationship episodes [14]. In this paper, we aim to address this gap by examining self-reported data on coital frequency and condom use during monogamous and concurrent relationship episodes from an

egocentric sexual network survey in three communities with high HIV prevalence around Cape Town, South Africa. Besides the concurrency status, we explore associations with a wide range of demographic and relationship characteristics, to identify other, potentially more important factors that influence coital frequency and condom use.

## Methods

### Study design and setting

We conducted a cross-sectional survey ( $n = 878$ ) from June 2011 to February 2012 in three urban disadvantaged communities in the greater Cape Town area to study associations between HIV status, sexual connectedness and age-disparity. The study design and protocol is explained in detail elsewhere [15]. In brief, the survey explored one-year sexual histories, with a focus on start and end dates of periods of sexual activity, age differences between sexual partners, sex frequency, condom use and the use of alcohol and recreational drugs. The questionnaire was administered in a safe and confidential mobile interview space, using audio computer-assisted self-interview (ACASI) technology on touch screen computers. ACASI has the benefit of providing privacy to participants and avoids the white coat effect when answering questions about sensitive topics. The ACASI featured a choice of languages and visual feedback of temporal information. All study communities participated in a previous TB/HIV surveillance study, from which HIV test results were anonymously linked to the survey dataset [16]. A list of participants from the TB/HIV surveillance study was generated for each of the three communities, and the names and associated addresses were randomly reordered. Field workers visited the homes of candidate survey participants in the order that they were placed on the list.

Of 1857 people randomly sampled from the TB/HIV surveillance study sampling frame, we were able to find 1115 (60.0% contact rate). For 197 people, the reason for non-retrieval after three attempts is unknown, while for, respectively, 511 and 34, relocation to an unknown new address and death were documented. Eighty-seven candidate participants were excluded, primarily due to visual or physical impairments that rendered participation in the study impossible. Of the remaining 1028, 878 (85.4% response rate) consented to participate.

### Participants and variables

Of the 878 survey respondents, 679 (77.3%) had at least one relationship in the last 12 months. These respondents reported on a total of 1567 relationship episodes from 1128 relationships. Relationship episodes with missing data for coital frequency ( $n = 193$ ), condom use ( $n = 5$ ), respondent age ( $n = 3$ ), partner age ( $n = 24$ ) respondent gender ( $n = 49$ ), race ( $n = 5$ ), completed education level ( $n = 1$ ) or employment status ( $n = 2$ ) were excluded. Furthermore, episodes were excluded if the respondents did not sleep with their partner in the past year ( $n = 14$ ) and if the reported ages of respondents were  $< 18$  years or  $> 70$  years ( $n = 42$ ). In the context of the South African HIV epidemic, the HIV prevalence is considerably higher in black and coloured communities than it is in other racial groups [17]. Our survey was

conducted in communities with high HIV prevalence, and consequently, very few people of Indian or white race were included in our sample. Therefore, 19 episodes from three respondents were excluded if the respondents were white, Indian or unknown race, leaving only episodes of black and coloured respondents in the analysis. The term *coloured* refers to a racial category in South Africa, and consists of racially mixed descendants of Europeans, indigenous populations and slaves from South and East Asia.

For up to five main partners and 15 casual partners, participants indicated the periods (episodes) they were in the relationship on a touch screen timeline [15]. A participant could select multiple different time periods for each partner. The dependent variables, frequency of intercourse and condom use, were asked for each episode indicated on the timeline. Periods of a week or longer during which participants indicated not having slept with a particular partner were counted as “breaks” between relationship episodes. For each relationship episode, participants were asked what the weekly average number of sex acts was (0, 1, 2, . . . 13, 14, 15, > 15) and how frequently they used condoms during sexual intercourse (always, sometimes, never). For each round of questions concerning a particular episode, the timing of the episode was highlighted on the touch screen timeline.

Figure 1 outlines how the concurrency status of each relationship episode was derived from the relationship history time line. Building on the defining characteristic of concurrency that individuals return to a previous partner (A) after having had intercourse with another partner (B), any episode for which this condition was true, was considered concurrent in the primary analysis [1]. Under this definition, as proposed by UNAIDS, 1A, 1B, 2A, 2B and 3B are concurrent episodes. However, this definition may be problematic as it lacks any indication of time scale over which the presence of overlap should be evaluated. Consequently, apparently very different kinds of “overlap” are grouped into the category of concurrent episodes, ranging from a situation in which participants move back and forth between sexual partners multiple times per week for many consecutive weeks, to a situation in which participants alternate

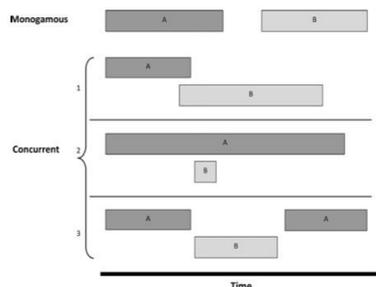


Figure 1. Schematic representation of monogamous and concurrent relationship episodes.

between multiple partners, but none of the episodes actually overlap (relationship type 3 in Figure 1). To explore how sensitive our results are to the definition of concurrency, we conducted two parallel analyses. In the first analysis, we applied the literal definition of concurrency according to the UNAIDS reference group (relation episodes 1A, 1B, 2A, 2B and 3B in Figure 1 defined as concurrent). In the second analysis, we only define episodes as concurrent if there is an actual temporal overlap of at least one week (3B in Figure 1 no longer included).

In addition to concurrency status, candidate explanatory variables for the variation in coital frequency and condom use included age ( $\leq 25/26-40/ >40$ ), race (coloured/black), religion (Christian/other religion/not religious), employment status (employed/unemployed), completed education level (none or primary/secondary/tertiary), age difference between partners (0-4/5-10/ >10), relationship duration ( $\leq 1$  week/2 weeks to 9 months/ >9 months) and partner type (casual/main).

#### Statistical analysis

First, the coital frequency and condom use data were tabulated and visualized by concurrency status and partner type, and descriptive summary statistics were calculated for all variables under investigation. Next, mixed effects logistic regression and mixed effects Poisson regression models were used to evaluate the effect of concurrency status, on consistent condom use and coital frequency, respectively. These models take into account the correlated nature of the data and variability in the data that stems from both inter- and intra-subject differences in repeated measurements (respondents may report on multiple relationships, which may each consist of multiple relationship episodes) [18]. Backward elimination procedures, based on likelihood ratio tests and Akaike's Information Criterion (AIC), were applied to assess whether employment status, completed education level, religion, age difference between partners, partner type and relationship duration were statistically independent correlates of coital frequency and consistent condom use, after adjusting for concurrency status, race, sex and age.

#### Ethical approval

The study was approved by the Stellenbosch University Health Research Ethics Committee (N11/03/093). Written, informed consent was obtained for each respondent prior to administration of the questionnaire.

#### Results

After exclusions, 1210 relationship episodes from 828 relationships reported by 527 sexually active respondents were retained. Tables 1 and 2 describe the demographic characteristics of these respondents and key attributes of their reported relationship episodes respectively.

The majority of respondents were black (80%) and female (69%). While females were clearly represented in higher numbers than males in our survey, the fraction of female respondents in our survey was not very different from that in the sampling frame (62%). Most respondents only reported one sexual partner in the last year (72%), and the vast majority of relationship episodes involved a main

**Table 1. Individual characteristics of participants in three urban Cape Town communities (aged 18–70 in 2011/2012)**

	n	%
<b>N</b>	<b>527</b>	
<b>Age</b>		
18–25 years	120	22.8
26–40 years	242	45.9
> 40 years	165	31.3
<b>Gender</b>		
Male	163	30.9
Female	364	69.1
<b>Race</b>		
Coloured	108	20.5
Black	419	79.5
<b>Education level</b>		
None or primary	162	30.7
Secondary	347	65.8
Tertiary	18	3.4
<b>Employment status</b>		
Employed	404	76.7
Unemployed	123	23.3
<b>Religion</b>		
Christian	350	66.4
Not religious	148	28.1
Other religion	29	5.5
<b>Numbers of partners last year</b>		
1	377	71.5
2	79	15.0
3	37	7.0
>3	27	5.1
<b>Casual partners last year</b>		
Yes	91	17.3
No	436	82.7

partner (82%). Forty-two percent (506/1210) of all episodes were concurrent according to the UNAIDS definition, while 41% (491/1210) were concurrent according to our modified definition. The median of the per-partner average coital frequency was two sex acts per week (IQR: 1–3; mean: 2.5), and consistent condom use (always used condoms) was reported in 36% of episodes. Only 28% (146/527) of the study sample reported consistent condom use in all episodes with all partners of the last year. Figures 2 and 3 depict average weekly coital frequency and condom use reported in each of the 1210 episodes, by concurrency status and partner type.

Figure 2 shows no immediately obvious, stark differences in coital frequencies in monogamous versus concurrent episodes. In the mixed effects regression analysis, presented in Table 3, there was no evidence for concurrency being associated with a lower average coital frequency. Rather, both definitions showed a slight, albeit non-significant, increase in coital frequency during concurrent episodes (UNAIDS definition:  $aIRR = 1.05$ ; 95% confidence interval (CI): 0.99–1.24

**Table 2. Attributes of relationship episodes from 520 participants in three urban Cape Town communities**

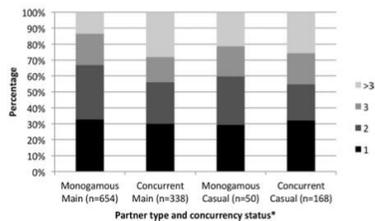
N	n	%
<b>Partner type</b>		
Main partner	992	82.0
Casual partner	218	18.0
<b>Concurrency status<sup>a</sup></b>		
Monogamous	704	58.2
Concurrent	506	41.8
<b>Concurrency status<sup>b</sup></b>		
Monogamous	719	59.4
Concurrent	491	40.6
<b>Condom use</b>		
Never	382	31.6
Sometimes	391	32.3
Always	437	36.1
<b>Duration</b>		
≤1 week	362	29.9
2 weeks to 9 months	490	40.5
>9 months	358	29.6
<b>Age difference between partners</b>		
<5 years	880	72.7
5–10 years	221	18.3
>10 years	109	9.0
<b>Average Coital frequency per episode</b>		
1	382	29.8
2	378	30.8
3	233	18.2
>3	217	21.2

<sup>a</sup>UNAIDS defined as any overlapping episode in which sexual intercourse with one partner occurs between two acts of intercourse with another partner. (Relationship episode types 1A, 1B, 2A, 2B, 3B from Figure 1.)

<sup>b</sup>Our modified definition of concurrency, which excludes relationship episode type 3B from Figure 1.

and modified definition: aIRR = 1.04; 95% CI: 0.98–1.23). Being female (aIRR = 0.83; 95% CI: 0.72–0.91), coloured (aIRR = 1.34; 95% CI: 1.13–1.48), obtaining a tertiary education (aIRR = 1.44; 95% CI: 1.12–1.96), having a relationship of 2 weeks to 9 months in duration (aIRR = 1.15; 95% CI: 1.05–1.30) and belonging to an “other” religion (aIRR = 1.27; 95% CI: 1.11–1.77) were shown to have a significant association with coital frequency in the model using the UNAIDS definition of concurrency. Using our modified concurrency definition did not qualitatively change these estimates.

The condom use outcomes shown in Figure 3 indicate higher consistent condom use in concurrent episodes with casual partners (55%; 44% in monogamous episodes), and similarly low levels of condom use in episodes with main partners, regardless of concurrency status (32–33%). In the mixed effects regression analysis, presented in Table 4,



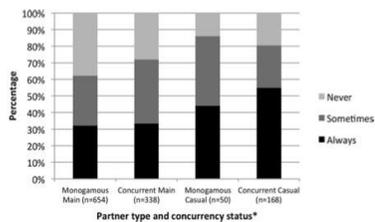
**Figure 2. Distribution of coital frequency, by partner type and concurrency status.**

\*Using the UNAIDS definition.

concurrency was not significantly associated with consistent condom use (UNAIDS definition: aOR = 1.01; 95% CI: 0.38–2.68 and modified definition: aOR = 1.48; 95% CI: 0.58–3.79), but race and relationship duration were. Being coloured (aOR = 0.08; 95% CI: 0.01–0.63) and having a relationship duration of more than 9 months (aOR = 0.08; 95% CI: 0.03–0.20) were associated with consistent condom use in the model using the UNAIDS definition of concurrency. Similarly to the coital frequency analysis, using our modified concurrency definition did not qualitatively change these estimates. Initial data exploration suggested that partner type was associated with consistent condom use as well, and that there might be effect modification of concurrency status by partner type and by gender. However, partner type could not be included in the final model because of quasi complete separation in the data tables. Furthermore, adding the interaction terms separately, did not improve model fit, and hence these interaction terms were not included in the final model.

## Discussion

Our findings have implications both for the debate around the role of concurrency in the spread of HIV, and more generally for priority setting in HIV prevention. The key factors that determine the role of concurrency in HIV transmission dynamics include: prevalence of concurrent relationships, duration of concurrent episodes, variability of HIV infectiousness with time since infection, connectedness of the entire sexual



**Figure 3. Distribution of condom use, by partner type and concurrency status.**

\*Using the UNAIDS definition.

**Table 3. Adjusted incident rate ratios for coital frequency using mixed effects models**

	UNAIDS concurrency definition <sup>a</sup>		Our modified concurrency definition <sup>b</sup>	
	aIRR	95% CI for aIRR	aIRR	95% CI for aIRR
<b>Concurrent</b>				
No	1.00		1.00	
Yes	1.05	0.99–1.24	1.04	0.98–1.23
<b>Age</b>				
18–25 years	1.00		1.00	
25–40 years	1.03	0.87–1.13	1.03	0.87–1.13
>40 years	0.98	0.81–1.11	0.98	0.81–1.11
<b>Gender</b>				
Male	1.00		1.00	
Female	0.83	0.72–0.91	0.81	0.72–0.91
<b>Race</b>				
Black	1.00		1.00	
Coloured	1.34	1.13–1.48	1.34	1.13–1.48
<b>Partner type</b>				
Main	1.00		1.00	
Casual	1.00	0.94–1.21	1.07	0.94–1.21
<b>Education</b>				
None or primary	1.00		1.00	
Secondary	1.12	0.94–1.22	1.12	0.94–1.22
Tertiary	1.44	1.12–1.96	1.45	1.12–1.96
<b>Duration</b>				
≤1 week	1.00		1.00	
2 weeks to 9 months	1.15	1.05–1.30	1.17	1.05–1.30
>9 months	1.07	0.95–1.22	1.08	0.95–1.23
<b>Religion</b>				
Christian	1.00		1.00	
Not religious	1.03	0.94–1.19	1.04	0.94–1.19
Other	1.27	1.11–1.77	1.28	1.11–1.77

aIRR, adjusted incident rate ratio; CI, confidence interval.

<sup>a</sup>Defined as any overlapping episode in which sexual intercourse with one partner occurs between two acts of intercourse with another partner. (Relationship episode types 1A, 1B, 2A, 2B, 3B from Figure 1.)

<sup>b</sup>Excludes relationship episode type 3B from Figure 1.

network and differences in frequencies of HIV exposures (unprotected sex acts) during monogamous versus concurrent episodes [2,10,12,19–21]. Given the large effect of sex frequency and consistent condom use on transmission risk, both coital dilution and increases in consistent condom use could substantially reduce the effect of concurrency on HIV transmission.

This study does not lend support to the coital dilution hypothesis, nor does it suggest increased condom use during periods of concurrency, after adjusting for confounding variables. Instead, in our study sample of black and coloured respondents from three communities around Cape

**Table 4. Adjusted odds ratios for consistent condom use using mixed effects models**

	UNAIDS concurrency definition <sup>a</sup>		Our modified concurrency definition <sup>b</sup>	
	aOR	95% CI for aOR	aOR	95% CI for aOR
<b>Concurrent</b>				
No	1.00		1.00	
Yes	1.01	0.38–2.68	1.48	0.58–3.79
<b>Age</b>				
18–25 years	1.00		1.00	
25–40 years	1.52	0.32–7.19	1.49	0.31–7.24
>40 years	0.87	0.15–4.94	0.90	0.15–5.31
<b>Gender</b>				
Male	1.00		1.00	
Female	1.07	0.27–4.21	1.14	0.28–4.59
<b>Race</b>				
Black	1.00		1.00	
Coloured	0.08	0.01–0.63	0.08	0.01–0.68
<b>Duration</b>				
≤1 week	1.00		1.00	
2 weeks to 9 months	0.46	0.21–1.01	0.42	0.18–0.95
>9 months	0.08	0.03–0.20	0.07	0.03–0.19

aOR, adjusted odds ratio; CI, confidence interval.

<sup>a</sup>Defined as any overlapping episode in which sexual intercourse with one partner occurs between two acts of intercourse with another partner. (Relationship episode types 1A, 1B, 2A, 2B, 3B from Figure 1.)

<sup>b</sup>Excludes relationship episode type 3B from Figure 1.

Town with high HIV prevalence, the coital frequency was higher, but not significantly higher, in concurrent compared to monogamous relationship episodes, regardless of the definition of concurrency. This finding is at odds with the survey findings from sub-Saharan Africa cited by Sawers *et al.* [4,11,22]. It is worth pointing out that Sawers *et al.* make incorrect inferences from Morris *et al.* [4] and Harrison *et al.* [22] by confusing and conflating concurrency status (monogamous versus concurrent) with relationship type (primary versus secondary). We believe the apparent discrepancies between these two studies cited by Sawers *et al.* and ours can be explained by differences in how coital frequency was measured and how concurrency status was assigned. In our survey, participants could indicate multiple relationship episodes with the same partner, with a resolution of one-week time blocks. This allowed us to observe relationships that consisted of multiple, disjointed episodes (14% of main relationships ( $n = 92$ ) and 13% of casual relationships ( $n = 22$ )) instead of one continuous time period.

Morris's categorization into "more frequent" and "less frequent" concurrent partners by design creates differences in sex frequency between different sexual partners. However, Morris's analysis does not confirm that coital frequencies are lower in concurrent versus monogamous relationships. Moreover, in the surveys reported by Morris *et al.* participants

were asked how many acts of sex they had over the course of the year for “primary” (more frequent) and “secondary” (less frequent) concurrent partnerships, assuming that these partnerships occurred as one continuous episode throughout the year with no gaps. Thus, the survey failed to take into account that some partnerships may have a low cumulative number of sex acts, but consist of one or many short episodes, during which the average coital frequency is high. In the same way, Harrison *et al.* failed to identify relationship episodes and measure coital frequency within each episode. Crucially, they did not restrict analysis of the time since last sex act with the last two sexual partners to respondents who were still in ongoing relationships with both these partners. Lack of knowledge of the concurrency status in this analysis of time since last sex act, precludes estimation of the effect of concurrency status on per-partner coital frequency. Sawers *et al.* may therefore have incorrectly inferred coital dilution from larger times since last sex with the second most recent partner.

To our knowledge, this is one of few studies that have attempted to identify behavioural and demographic correlates of coital frequency in concurrent and monogamous relationships [2,23–25]. In our study sample, being coloured, male and having a tertiary education; being in a relationship for a period of 2 weeks to 9 months; and belong to an “other” religion were independent, individual-level predictors of higher coital frequency.

Our crude estimators for consistent condom use in monogamous and concurrent relationship episodes (Figure 3) compare well with related statistics previously reported. In a survey among young black people around Cape Town, 44% of men with a history of concurrency reported consistent condom use [26]. Further, Chopra *et al.* reported more consistent condom use with casual partners than with “steady” partners in a cohort of young Cape Town men of whom 98% reported concurrent relationships in the last three months [27]. Similarly, Maher *et al.* observed that condom use with concurrent partners was more frequent if partnerships were casual instead of “regular”, non-spousal [7].

Results from the mixed effects regression analysis do not provide evidence for increased condom use during concurrency. Other studies, however, have found significant associations between condom use and concurrency status. Of note, Steffenson *et al.* found that in South African men and women aged 15–24, those who had at least one concurrent relationship in the last year (“concurrents”) used condoms less frequently than people in monogamous relationships (“monogamists”) [28]. The discrepancy between their study results and ours might be accounted for by the fact that our analysis was done at the level of relationship episodes, and compares all of the monogamous to all of the concurrent episodes, while adjusting for a range of confounding variables. In contrast, Steffenson *et al.* measured concurrency status at the level of an individual and then compared condom use during only the most recent relationship in “concurrents” and “monogamists”. They, therefore, were not able to accurately determine if concurrent relationships, much less concurrent episodes, are associated with less consistent condom use. Another study, conducted by Kasamba *et al.*, explored condom use in spousal and extra-

spousal partnerships and found that men who had extra-spousal partnerships were more likely to have ever used condoms with their spouse [29]. Direct comparison with our findings is limited by the fact that they measured “ever having used condoms” and classified relationships into spousal and extra-spousal relationships. We measured “always used a condom” rather than “ever used a condom” because it is a more meaningful indicator of HIV risk aversion.

Implications of our findings for HIV prevention efforts follow primarily from the observation that consistent condom use was generally low, especially in relationships with main partners. Consistent condom use is known to be extremely hard to achieve in long-term, trusting relationships [30], even if they involve transactional sex [31]. Although consistent condom use was more frequently reported with casual partners, as was also seen elsewhere [32,33], there is still a lot of potential for averting HIV transmissions in casual relationships, especially since casual partners may carry a higher burden of sexually transmitted infections, which are known to facilitate HIV transmission [34–37].

Our study has four main limitations. First, in our study, respondents could only report one average weekly coital frequency per episode, regardless of the episode’s duration. Consequently, this self-reported average would only be affected minimally, if at all, if coital frequency was temporarily lower during times of concurrency with an episode that overlapped the index episode for a small fraction. Second, left and right censoring of relationships may have led to misclassification of some episodes as monogamous because we had no knowledge of future episodes and episodes that took place more than a year before the survey. Third, the candidate individual-level predictor variables (i.e. religion, employment status, education level, age, sex and race) we explored were asked only at the time of the survey, but used to predict past behaviour (i.e. coital frequency and condom use). Theoretically, these variables may not have stayed constant over the one-year relationship history window. Lastly, our survey data may be subject to bias due to possible dependent errors in reporting concurrency, coital frequency and condom use. We do note, however, that this bias may also have been present in the egocentric survey data that was cited by Sawers *et al.* to support the coital dilution hypothesis. Hence, this bias alone cannot explain the difference between our observations and those cited previously in support of coital dilution.

Despite these limitations, our study had several strengths, which we believe support the accuracy of our results. Rather than face-to-face interviewing, the survey was conducted using ACASI. While comparisons of ACASI and more traditional survey methods have been mixed, several studies that compared ACASI methods with face-to-face interviews in the African context have indicated that participants are more likely to report sexual risk behaviours while using ACASI [38–42]. In addition, we have performed a dedicated analysis of the user-friendliness, privacy and truthfulness of our ACASI instrument. The key conclusion of this paper is that most participants in our survey found the ACASI modality to be acceptable, private, and user-friendly. Moreover, our results indicate less social desirability bias when reporting on

multiple, concurrent partners, than in the face-to-face interviews used in Demographic and Health Surveys done in Southern Africa [43]. Furthermore, respondents were asked to place the episodes for each of their relationships in the past year directly on a timeline, progressively from the oldest to the most recent relationship. Thus, the timeline and the episodes of earlier relationships provided visual reference points, which facilitated internal consistency of a respondent's relationship history [44,45]. Finally, our study is unique in that it allowed participants to define their relationships as a series of episodes, which more accurately portrays how people engage in relationships. In reality, relationships are not always continuous: they often have periods of sexual activity and inactivity, and sexual behaviours may not be the same for each new period of a relationship.

### Conclusions

In conclusion, we found no evidence for coital dilution, i.e. for a decreased per-partner sex frequency, or for increased condom use during concurrent relationship episodes in three communities around Cape Town with high HIV prevalence, after adjusting for confounding variables. Instead, concurrency was associated with a slight, borderline-significant (at  $\alpha = 0.05$ ) increase in coital frequency. The main implication of our findings for the concurrency debate is that, if the frequency of unprotected sex with each of the sexual partners is sustained during concurrent relationships, HIV-positive individuals with concurrent partners may disproportionately contribute to onward HIV transmission. Additional analyses from other geographic and epidemiological settings are needed to create a larger body of evidence related to coital frequency and condom use in monogamous and concurrent relationship episodes, and more generally, to deepen our understanding of the determinants of coital frequency and consistent condom use.

### Authors' affiliations

<sup>1</sup>The South African Department of Science and Technology/National Research Foundation (DST/NRF), Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa; <sup>2</sup>International Centre for Reproductive Health, Ghent University, Ghent, Belgium; <sup>3</sup>Center for Statistics, Hasselt University, Diepenbeek, Belgium; <sup>4</sup>Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium; <sup>5</sup>Centre for Health Economic Research and Modelling Infectious Diseases, Vaccine and Infectious Disease Institute, University of Antwerp, Wilrijk, Belgium

### Competing interests

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

### Authors' contributions

WD, ND, RB, MT, AW and NH jointly designed the ACASI survey. RB coordinated data collection. FM and RB conducted the statistical analysis. NH supervised the data analysis. WD and RB wrote the first draft manuscript. All authors contributed during the editing process and approved the final, submitted manuscript.

### Acknowledgements

The authors wish to thank the survey respondents for their participation in this study. We are also indebted to the research assistants for countless hours of dedicated fieldwork and to the Desmond Tutu TB Centre at Stellenbosch University for their guidance in coordination and planning of this cross-sectional sexual behaviour survey.

**Funding sources:** The Flemish Interuniversity Council (VLIR) and the Flemish Scientific Research Fund (FWO) provided financial support for the survey.

WD, RB and AW received funding from the Canadian International Development Agency (CIDA). FM was funded by FWO, and NH acknowledges support from the University of Antwerp scientific chair in Evidence-Based Vaccinology, financed in 2009–2012 by an unrestricted donation from Pfizer.

### References

- UNAIDS Reference Group on Estimates Modelling and Projections: Working Group on Measuring Concurrent Sexual Partnerships. HIV: consensus indicators are needed for concurrency. *Lancet*. 2010;375(9715):621–2.
- Goodreau SM, Cassels S, Kasprzyk D, Montano DE, Greek A, Morris M. Concurrent partnerships, acute infection and HIV epidemic dynamics among young adults in Zimbabwe. *AIDS Behav*. 2012;16(2):312–22.
- Mah TL, Shelton JD. Concurrency revisited: increasing and compelling epidemiological evidence. *J Int AIDS Soc*. 2011;14:33.
- Morris M, Epstein H, Wawer M. Timing is everything: international variations in historical sexual partnership concurrency and HIV prevalence. *PLoS One*. 2010;5(11):e14092.
- Lurie MN, Rosenthal S. Concurrent partnerships as a driver of the HIV epidemic in sub-Saharan Africa? The evidence is limited. *AIDS Behav*. 2010;14:17–24.
- Lagarde E, Auvert B, Carael M, Lacourou M, Ferry B, Akam E, et al. Concurrent sexual partnerships and HIV prevalence in five urban communities of sub-Saharan Africa. *AIDS*. 2001;15(7):877–84.
- Malher D, Waswa L, Karabarinde A, Baisley K. Concurrent sexual partnerships and associated factors: a cross-sectional population-based survey in a rural community in Africa with a generalised HIV epidemic. *BMC Public Health*. 2011;11:651.
- Mishra V, Binami-Van Assche S. Concurrent sexual partnerships and HIV infection: evidence from national population-based surveys. DHS Working papers. Calverton, Maryland: Macro International Inc; 2009.
- Tanser F, Barnighausen T, Hund L, Garnett GP, McGrath N, Newell ML. Effect of concurrent sexual partnerships on rate of new HIV infections in a high-prevalence, rural South African population: a cohort study. *Lancet*. 2011;378(9787):247–55.
- Sawers L, Stillwaggon E. Concurrent sexual partnerships do not explain the HIV epidemics in Africa: a systematic review of the evidence. *J Int AIDS Soc*. 2010;13:34.
- Sawers L, Isaac AG, Stillwaggon E. HIV and concurrent sexual partnerships: modelling the role of coital dilution. *J Int AIDS Soc*. 2011;14:44.
- Bolly MC, Alary M, Baggaley RF. Neglected issues and hypotheses regarding the impact of sexual concurrency on HIV and sexually transmitted infections. *AIDS Behav*. 2012;16(2):304–11.
- Reniers G, Watkins S. Polygyny and the spread of HIV in sub-Saharan Africa: a case of benign concurrency. *AIDS*. 2010;24(2):299–307.
- Epstein H, Swidler A, Gray R, Reniers G, Parker W, Parkhurst J, et al. Measuring concurrent partnerships. *Lancet*. 2010;375:1869.
- Delva W, Beauclair R, Welte A, Vansteelandt S, Hens N, Aerts M, et al. Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol. *BMC Public Health*. 2011;11:616.
- Ayles HM, Sismanidis C, Beyers N, Hayes RJ, Godfrey-Faussett P. ZAMSTAR, The Zambia South Africa TB and HIV Reduction study: design of a 2 x 2 factorial community randomized trial. *Trials*. 2008;9:63.
- Shisana O, Simbayi LC, Rehle T, Zungu NP, Zuma K, Ngogo N, et al. South African National HIV Prevalence, Incidence, Behaviour and Communication Survey, 2008: the health of our children. Cape Town: HSRC Press; 2010.
- Cnaan A, Laird NM, Slator P. Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Stat Med*. 1997;16(20):2349–80.
- Eaton JW, Hallett TB, Garnett GP. Concurrent sexual partnerships and primary HIV infection: a critical interaction. *AIDS Behav*. 2011;15(4):687–92.
- Powers KA, Hoffman IF, Ghani AC, Hossenlopp MC, Pilcher CD, Price MA, et al. Sexual partnership patterns in Malawi: Implications for HIV/STI transmission. *Sex Transm Dis*. [Research Support, N.I.H., Extramural]. 2011;38(7):657–66.
- Kretzschmar M, White RG, Carael M. Concurrency is more complex than it seems. *AIDS*. 2009;24:313–5.
- Harrison A, Cleland J, Frohlich J. Young people's sexual partnerships in KwaZulu-Natal, South Africa: patterns, contextual influences, and HIV risk. *Stud Fam Plann*. 2008;39(4):295–308.
- Westoff CF. Recent trends in rates of sexual activity in sub-Saharan Africa. *J Biosoc Sci*. 2007;39(6):895–904.

24. Reniers G, Tlailly R. Polygyny, partnership concurrency, and HIV transmission in sub-Saharan Africa. *Demography*. 2012;49:1075–101.
25. Xu H, Luke N, Zulu EM. Concurrent sexual partnerships among youth in urban Kenya: prevalence and partnership effects. *Popul Stud*. 2010;64(3):247–61.
26. Maughan-Brown B. Variation in concurrent sexual partnerships and sexually transmitted diseases among African men in cape town, South Africa. *Sex Transm Dis*. 2012;39(7):537–42.
27. Chopra M, Townsend L, Johnston L, Mathews C, Tomlinson M, O'Bra H, et al. Estimating HIV prevalence and risk behaviors among high-risk heterosexual men with multiple sex partners: use of respondent-driven sampling. *J Acquir Immune Defic Syndr*. 2009;51(1):72–7.
28. Steffenson AE, Pettifor AE, Seage GR 3rd, Rees HV, Cleary PD. Concurrent sexual partnerships and human immunodeficiency virus risk among South African youth. *Sex Transm Dis*. 2011;38(6):459–66.
29. Kasamba I, Sully E, Weiss HA, Baisley K, Maher D. Extraspousal partnerships in a community in rural Uganda with high HIV prevalence: a cross-sectional population-based study using linked spousal data. *J Acquir Immune Defic Syndr*. 2011;58(1):108–14.
30. Chimbiri AM. The condom is an 'intruder' in marriage: evidence from rural Malawi. *Soc Sci Med*. 2007;64(5):1102–15.
31. Voeten HA, Egesh OB, Varkevisser CM, Habbema JD. Female sex workers and unsafe sex in urban and rural Nyanza, Kenya: regular partners may contribute more to HIV transmission than clients. *Trop Med Int Health*. 2007;12(2):174–82.
32. Chimbindi NZ, McGrath N, Herbst K, San Tint K, Newell ML. Socio-demographic determinants of condom use among sexually active young adults in rural KwaZulu-Natal, South Africa. *Open AIDS J*. 2010;4:88–95.
33. Westercamp N, Mattson CL, Madonia M, Moses S, Agot K, Ndinya-Achola JO, et al. Determinants of consistent condom use vary by partner type among young men in Kisumu, Kenya: a multi-level data analysis. *AIDS Behav*. 2010;14(4):949–59.
34. Onoya D, Reddy P, Sifunda S, Lang D, Wingood GM, van den Borne B, et al. Transactional sexual relationships, sexually transmitted infection risk, and condom use among young Black Women in peri-urban areas of the Western Cape Province of South Africa. *Womens Health Issues*. 2012;22(3):e277–82.
35. Wasserheit JN. Epidemiological synergy. Interrelationships between human immunodeficiency virus infection and other sexually transmitted diseases. *Sex Transm Dis*. 1992;19(2):61–77.
36. Fleming DT, Wasserheit JN. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. *Sex Transm Infect*. 1999;75(1):3–17.
37. Mahiane SG, Legeai C, Taljaard D, Latouche A, Puren A, Peillon A, et al. Transmission probabilities of HIV and herpes simplex virus type 2, effect of male circumcision and interaction: a longitudinal study in a township of South Africa. *AIDS*. 2009;23(3):377–83.
38. Dolezal C, Marhefka SL, Santamaria EK, Leu CS, Brackis-Cott E, Mellins CA. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Arch Sex Behav*. 2012;41(2):401–10.
39. Hewett P, Mensch B, Erulcar A. Consistency in the reporting of sexual behavior by adolescent girls in Kenya: a comparison of interviewing methods. *Sex Transm Infect*. 2004;80(Suppl 2):i43–8.
40. Kissinger P, Rice J, Farley T, Trim S, Jewitt K, Margava V, et al. Application of computer-assisted interviews to sexual behavior research. *Am J Epidemiol*. 1999;149(10):950–4.
41. Phillips AE, Gomez GB, Boily MC, Garnett GP. A systematic review and meta-analysis of quantitative interviewing tools to investigate self-reported HIV and STI associated behaviours in low- and middle-income countries. *Int J Epidemiol*. 2010;39(6):1541–55.
42. van der Elst E, Okuku H, Nakamya P, Muhaari A, Davies A, McClelland R, et al. Is audio computer-assisted self-interview (ACASI) useful in risk behavior assessment of female and male sex workers, Mombasa, Kenya? *PLoS One*. 2009;4(5):e5340.
43. Beauclair R, Meng F, Deprez N, Temmerman M, Welte A, Hens N, et al. Evaluating audio computer assisted self-interviews in urban South African communities: evidence for good suitability and reduced social desirability bias of a cross-sectional survey on sexual behaviour. *BMC Med Res Methodol*. 2013;13:11.
44. Kabiru CW, Luke N, Izugbara CO, Zulu E. The correlates of HIV testing and impacts on sexual behaviour: evidence from a life history study of young people in Kisumu, Kenya. *BMC Public Health*. 2010;10:412.
45. Luke N, Clark S, Zulu E. The relationship history calendar: improving the scope and quality of data on youth sexual behavior. *Demography*. 2011;48(3):1151–76.

# APPENDIX D

Beauclair R et al. *Journal of the International AIDS Society* 2015, **18**:19372  
<http://www.jiasociety.org/Index.php/JIAS/article/view/19372> | <http://dx.doi.org/10.7448/IAS.18.1.19372>



## Research article

### Concurrent partnerships in Cape Town, South Africa: race and sex differences in prevalence and duration of overlap

Roxanne Beauclair<sup>5,1,2</sup>, Niel Hens<sup>3,4</sup> and Wim Delva<sup>1,2,3</sup>

<sup>5</sup>Corresponding author: Roxanne Beauclair, c/o StIAS, Private Bag X1, Matieland, 7602 Stellenbosch, South Africa. Tel: +27 21 808 2589, Fax: +27 21 808 2586. ([roxanne.beauclair@gmail.com](mailto:roxanne.beauclair@gmail.com))

#### Abstract

**Introduction:** Concurrent partnerships (CPs) have been suggested as a risk factor for transmitting HIV, but their impact on the epidemic depends upon how prevalent they are in populations, the average number of CPs an individual has and the length of time they overlap. However, estimates of prevalence of CPs in Southern Africa vary widely, and the duration of overlap in these relationships is poorly documented. We aim to characterize concurrency in a more accurate and complete manner, using data from three disadvantaged communities of Cape Town, South Africa.

**Methods:** We conducted a sexual behaviour survey ( $n = 878$ ) from June 2011 to February 2012 in Cape Town, using Audio Computer-Assisted Self-Interviewing to collect sexual relationship histories on partners in the past year. Using the beginning and end dates for the partnerships, we calculated the point prevalence, the cumulative prevalence and the incidence rate of CPs, as well as the duration of overlap for relationships begun in the previous year. Linear and binomial regression models were used to quantify race (black vs. coloured) and sex differences in the duration of overlap and relative risk of having CPs in the past year.

**Results:** The overall point prevalence of CPs six months before the survey was 8.4%: 13.4% for black men, 1.9% for coloured men, 7.8% black women and 5.6% for coloured women. The median duration of overlap in CPs was 7.5 weeks. Women had less risk of CPs in the previous year than men (RR 0.43; 95% CI: 0.32–0.57) and black participants were more at risk than coloured participants (RR 1.86; 95% CI: 1.17–2.97).

**Conclusions:** Our results indicate that in this population the prevalence of CPs is relatively high and is characterized by overlaps of long duration, implying there may be opportunities for HIV to be transmitted to concurrent partners.

**Keywords:** concurrent partnerships; sexual concurrency; HIV prevention; South Africa; sexual behaviour and HIV; sexual risk behaviour.

To access the supplementary material to this article please see Supplementary Files under Article Tools online.

Received 5 August 2014; Revised 19 November 2014; Accepted 13 January 2015; Published 18 February 2015

Copyright: © 2015 Beauclair R et al; licensee International AIDS Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

#### Introduction

Concurrent partnerships (CPs) have been identified as a potentially important facilitator of the HIV epidemic in Southern Africa [1–5]. CPs have been defined as “overlapping sexual partnerships in which sexual intercourse with one partner occurs between two acts of intercourse with another partner” [6]. Theoretically, CPs facilitate faster spread of HIV because of two different effects at the sexual network level: 1) earlier partnerships begun by the index partner are later exposed to any infections transmitted by an additional partner and 2) the time to secondary transmission is shortened because the infected person does not need to terminate one partnership before starting another. Proponents of this hypothesis have demonstrated through modelling studies that the effect of CPs on HIV transmission is large when high infectiousness during the acute stage of infection is taken

into account [1,3,5]. Evidence from ecological studies has also shown that prevalence of HIV is correlated with prevalence of CPs [2,4,7,8].

In contrast with that evidence, data from rural KwaZulu-Natal, South Africa, indicated that women living in communities with high prevalence of CPs among men did not have increased risk of acquiring HIV [9]. Furthermore, a randomized controlled trial in Kisumu, Kenya, found that being infected with HIV at baseline was not associated with having CPs [10]. Reniers and Watkins proposed the coital dilution hypothesis as an explanation for why CPs are not likely to drive the spread of HIV [11]. They assert that as individuals acquire additional CPs, the average sex frequency per partner drops. Therefore, if the total number of partnerships in a population remains constant and the average coital frequency in those partnerships falls, then the total number of sex acts

in the population must fall and result in a slower spread of HIV [11,12]. However, support for this theory has been mixed [13,14].

Go *et al.* argue that in order for CPs to be a major driver of HIV epidemics, three requirements must be met: first, the prevalence of CPs in a population must be high; second, the number of CPs should be moderately high for the average individual; and third, the duration of overlap should be relatively long [15]. It is currently difficult to tell if these conditions have been met in studied populations because definitions and measurement of CPs to date have been problematic. Many studies define CPs as any report of extra-spousal relationships (or casual relationships if they report a main partner) in the preceding 12 months [2,4,16,17]. These definitions assume that all participants are having sex with their spousal/main partners before and after the extra-spousal/casual sex. Moreover, this “direct” approach to measuring and defining concurrency precludes the capturing of beginning and end dates, which has meant that some studies could not calculate point prevalence of CPs, nor the duration of relationships, much less the length of time the relationships were overlapping [18–21]. In other studies where the duration of relationships is known, the measurement of relationship dates was accurate only up to one month [13,22]. This means that those studies could only capture relationships that overlapped by one month or more, missing potentially many more short-term concurrent relationships in their prevalence estimates. Still, some studies use better definitions of CPs but questions about CPs are asked in face-to-face interviews [9,23,24], which may promote considerable social desirability bias [25] and underestimate the true prevalence of CPs.

There have been calls for a more refined definition of CPs and guidelines for measuring it [26]. The UNAIDS-proposed indicator of CPs is the point prevalence of having more than one sexual partnership, six months before the interview [6]. We believe this indicator may only be part of the solution. Even if point prevalence of CPs is calculated, further steps to improve accuracy must be taken, as well as combining this indicator with other measures of concurrency in order to produce a more complete and useful picture of CP dynamics. Not only is a more complete description of CPs important for an intuitive understanding of how HIV spreads in a population but this would also allow for better parameterization of epidemiological models that are used to study HIV (combination) prevention and HIV transmission dynamics.

In South Africa, where there are an estimated 6.4 million people currently living with HIV [27], it is important to obtain accurate estimates of CP point prevalence and the duration of overlapping relationships in order to understand why the prevalence of HIV is so high. Furthermore, race and sex are primary determinants of social, health, economic and educational opportunities in South Africa [28], and thus the relationship between them and CPs needs to be evaluated to determine which groups should be targeted for potential CPs and HIV risk-reduction interventions. There have been studies in Cape Town, South Africa, that have examined how race and sex are related to CPs [18,29–31]; however, none investigated the duration of overlap or incidence of entering a CP.

With this in mind, we analyzed data from a cross-sectional survey conducted in urban communities of Cape Town with a high prevalence of HIV. The survey questions were administered using Audio Computer-Assisted Self-Interviewing (ACASI), so as to minimize social desirability bias. We characterized concurrency in this population by estimating the point prevalence, cumulative prevalence, incidence and degree distribution of CPs. We also described the duration of overlaps for relationships begun in the previous year and the relative risk of having CPs for different race and sex groups.

## Methods

### Study design and setting

We conducted a cross-sectional survey ( $n = 878$ ) from June 2011 to February 2012 in three urban communities of Cape Town with high prevalence of HIV. The study communities were characterized by high unemployment, informal housing and very little post-secondary education [32]. These communities primarily have residents that identify as black or coloured. The participants were randomly sampled from a previous community randomized trial that aimed to reduce the prevalence of TB and HIV using novel public health interventions, the features of which have been previously published [33].

The survey was administered in a mobile office space on touchscreen computers. ACASI was used to provide participants with privacy while answering sensitive questions about their sexual behaviours. Participants could indicate on a visual timeline the beginning and end dates of relationships for up to five main partners and ten casual partners that they had in the previous 12 months. The dates were accurate up to one week within a month. Additional questions about condom use, age of partners, alcohol and drug use at first sex, sex frequency and proximity to partners were asked for relationships. The study was approved by the Stellenbosch University Health Research Ethics Committee (N11/03/093) and written, informed consent was obtained from each participant. Further details of our sampling strategy and study design have been published elsewhere [14,25,34].

### Participants

The survey had a contact rate of 60% ( $n = 1115$ ) for individuals enumerated in the sample frame, and a response rate of 85.4% ( $n = 878$ ) for those we contacted. We excluded participants from the analysis who were not between the ages of 15 and 70, or who had missing ages ( $n = 34$ ); had missing sex information ( $n = 30$ ); were not heterosexual ( $n = 50$ ); and who did not identify as black or coloured race, or had missing race information ( $n = 14$ ). We chose to include only the racial groups that have been previously characterized as having high HIV prevalence in their populations. The term “coloured” is an official racial category in South Africa, and is used to describe people who have racially mixed ancestry from Europe, indigenous populations, and South and East Asia. After exclusions, 750 participants remained and they reported on 1003 relationships. Of the 750 remaining participants, 148 did not report on any partner in the previous year. Characteristics of those participants can be found in the Supplementary file.

### Statistical analysis

All statistical analyses were performed using Stata statistical software, version 12.0 (Stata-Corp Inc., College Station, TX, USA). First, we calculated frequency distributions and summary statistics of relationship characteristics for participants ( $n = 602$ ) that reported having had at least one relationship in the past year ( $n = 1003$  relationships). Some relationship questions were only asked for relationships that started in the previous 12 months ( $n = 415$  relationships).

Next, by sex and race, we estimated the one-year cumulative prevalence, point prevalence six months before the survey and incidence rate of CPs. We calculated these in two ways: 1) all participants were included in the denominator ( $n = 750$ ), which is the indicator recommended by UNAIDS [6] and 2) only participants who reported having sex in the past year were in the denominator ( $n = 528$ ). Two relationships were categorized as concurrent if the beginning date of one relationship overlapped with the end date of the other relationship by at least one week. We expressed the incidence rate as number of cases per 1000 person-years, and calculated it by dividing the total number of CPs that occurred in the previous year by the total person-years spent at risk and then multiplying by 1000. Each person contributed one year of person-time to the denominator, because each day of the previous year provided another opportunity for a person to begin a new concurrent relationship. Central to our calculation of the incidence of CPs is the notion that a “case” is the occurrence of a concurrent relationship, not the event of an individual entering a new relationship while already engaged in an earlier relationship.

Next, for sexually active participants, we computed the median and range of the total number of partners in the previous year (regardless of whether they were concurrent or serially monogamous) and the total number of CPs. These same data are also presented as a frequency distribution for total numbers of partners ( $1/2 / > = 3$ ) and CPs ( $0/2 / > = 3$ ) in the past year. It is important to note here that it is not possible to have only one CP. At the moment when one of the relationships begins to overlap with another, they both become CPs, and thus the number of CPs will equal two. We calculated average duration of overlap for each relationship because each relationship could potentially overlap with more than one other relationship. Therefore, we present the median of participants’ average duration of overlaps, by sex and race. In this calculation, we excluded relationships that began outside of the one-year window because we do not have information on their complete durations.

In addition, we constructed univariate and multivariable binomial regression models with a log link function, in order to examine the risk of having a CP in the past 12 months for participants in different race-sex groups. Finally, we used simple and multiple linear regression models to look at the effect of race and sex on the duration of overlap for CPs begun in the previous year. Here, the unit of analysis was the relationship, and because each participant could report on more than one relationship, we used clustered sandwich estimators. For this sub-analysis, we used only relationships that began in the previous year ( $n = 301$ ), because we did not have complete durations for relationships that began

outside of the one-year window. For each of the above regression analyses, a multivariable model was fit that included a race–sex interaction term, in order to see if race modifies the effect of sex.

### Results

Among participants who reported at least one relationship in the previous year, 20.9% ( $n = 126$ ) were coloured and 79.1% ( $n = 476$ ) were black; 32.4% ( $n = 195$ ) were male and 67.6% ( $n = 407$ ) were female. Characteristics of their relationships are presented in Table 1. Overall, most relationships reported by participants were from self-described “main” partnerships (76.1%) and the age difference between female and male partners tended to be less than five years (57.3%). Of note, black females had the highest percentage of age-disparate relationships (41.5%) and this demographic group had the highest proportion believing that their partner had other partners (58.6%). Coloured men had the highest proportion of casual relationships (42.3%) and non-condom use at the last sex (59.4%). Coloured women reported the highest percentage of relationships where the partner lived in the same home (34.6%) and the lowest fraction of using a condom at last sex (26.9%). Black men had the lowest median age of partner (26 years).

Estimates of prevalence, incidence rates and median duration of CPs overlap are presented in Table 2. Among all participants, the point prevalence of CPs was 8.4% ( $n = 63$ ), the one-year cumulative prevalence was 18.4% ( $n = 138$ ) and the incidence rate was 557.3 concurrent partners per 1000 person-years. The median duration of overlap for participants with CPs begun in the previous year was 7.5 weeks (IQR: 2.2–24), with coloured women having the longest duration of overlap at 26.5 weeks (IQR: 16–32). Black men had the highest point prevalence (13.4%), cumulative prevalence (34.2%) and incidence (1117.7 per 1000 person-years), compared to the other subgroups. The maximum number of relationships anyone had in the previous year was 12, with 29.5% of sexually active participants having more than one relationship. Overall, 11.7% of sexually active participants had three or more CPs in the previous year. Men of both races tended to have more CPs than women, and for both sexes, black participants had more CPs than coloured participants. For both black and coloured men, there was a higher frequency of three or more CPs (25.9 and 12.1%, respectively) than just two CPs (20.0 and 9.1%, respectively). Overall, of the 156 participants who had two or more relationships in the past year, 85.9% ( $n = 134$ ) had concurrent as opposed to serially monogamous partnerships. In all of the race–sex subgroups except coloured women, the cumulative prevalence of CPs was substantially higher than the estimated point prevalence. In particular, among coloured men, the cumulative prevalence (13.0%) was nearly seven times higher than the point prevalence (1.9%). This difference is most likely because of the relatively high incidence of CPs (500 per 1000 person-years) and short median duration of overlaps (four weeks).

The results from the binomial regression analysis can be found in Table 3. Both the crude and adjusted risk ratios indicate that women were at a lower risk than men of having had CPs in the previous 12 months (aRR 0.43; 95% CI: 0.32–0.57). In the adjusted model, black people were almost twice

**Table 1. Relationship characteristics of participants by race and sex**

	Relationships ( <i>n</i> = 1003)				
	Men		Women		
	Black <i>n</i> = 362	Coloured <i>n</i> = 71	Black <i>n</i> = 458	Coloured <i>n</i> = 112	
Age of partner, median (IQR)	32 (25–42)	26 (21–37)	37 (28–49)	34 (27–44)	34 (27–45)
Age disparate, <i>n</i> (%)					
Non-age disparate	575 (57.3)	201 (55.5)	48 (67.6)	250 (54.6)	76 (67.9)
Age disparate <sup>a</sup>	353 (35.2)	118 (32.6)	16 (22.5)	190 (41.5)	29 (25.9)
Missing	75 (7.5)	43 (11.9)	7 (9.9)	18 (3.9)	7 (6.3)
Partner type, <i>n</i> (%)					
Main	763 (76.1)	226 (62.4)	41 (57.8)	407 (88.9)	89 (79.5)
Casual	240 (23.9)	136 (37.6)	30 (42.3)	51 (11.1)	23 (20.5)
Still ongoing, <i>n</i> (%)					
Yes	617 (61.5)	173 (47.8)	33 (46.5)	323 (70.5)	88 (78.6)
No	364 (36.3)	181 (50.0)	33 (46.5)	126 (27.5)	24 (21.4)
I do not know	22 (2.2)	8 (2.2)	5 (7.0)	9 (2.0)	0 (0.0)
	Relationships started in the past year ( <i>n</i> = 415)				
	Men		Women		
	Black <i>n</i> = 217	Coloured <i>n</i> = 32	Black <i>n</i> = 140	Coloured <i>n</i> = 26	
Where does your partner live in proximity to you?, <i>n</i> (%)					
Same house	68 (16.4)	30 (13.8)	7 (21.9)	22 (15.7)	9 (34.6)
Different house, same township	178 (42.9)	99 (45.6)	14 (43.8)	62 (44.3)	3 (11.5)
Different township, Western Cape	88 (21.2)	40 (18.4)	8 (25.0)	33 (23.6)	7 (26.9)
Different province, South Africa	40 (9.6)	23 (10.6)	1 (3.1)	13 (9.3)	3 (11.5)
I do not know/missing	41 (9.9)	25 (11.5)	2 (6.3)	10 (7.1)	4 (15.4)
Partner had other partners, <i>n</i> (%)					
Yes	188 (45.3)	88 (40.6)	14 (43.8)	82 (58.6)	4 (15.4)
No	168 (40.5)	90 (41.5)	16 (50.0)	44 (31.4)	18 (69.2)
I do not know/missing	59 (14.2)	39 (18.0)	2 (6.25)	14 (10.0)	4 (15.4)
Condom at last sex with partner, <i>n</i> (%)					
Yes	195 (47.0)	105 (48.4)	11 (34.4)	72 (51.4)	7 (26.9)
No	180 (43.4)	88 (40.6)	19 (59.4)	58 (41.4)	15 (57.7)
I do not know/missing	40 (9.6)	24 (11.1)	2 (6.3)	10 (7.1)	4 (15.4)
On alcohol at first sex, <i>n</i> (%)					
Yes	111 (26.8)	69 (31.8)	12 (37.5)	23 (16.4)	7 (26.9)
No	265 (63.9)	126 (58.1)	18 (56.3)	106 (75.7)	15 (57.7)
I do not know/missing	39 (9.4)	22 (10.1)	2 (6.3)	11 (7.9)	4 (15.4)

<sup>a</sup>Greater than five years difference between the female partner and the male partner.  
 IQR, inter-quartile range.

as likely to have had CPs (aRR 1.86; 95% CI: 1.17–2.97). The interaction term added to the multivariable model was not significant. The results in Table 4 show the effect of sex and race on the duration of overlap for 301 concurrent relationships belonging to 133 people. There is a significant effect of sex and race on duration of overlap when an interaction term is included. In coloured participants, the average duration of overlap was 19.43 weeks longer for females than males (95%

CI: 7.09–31.76). The average duration of overlap for relationships begun in the previous year was 8.1 weeks longer for black men compared to coloured men (95% CI: 2.6–13.6).

### Discussion

Using data from our sexual behaviour survey located in disadvantaged communities of Cape Town, we aimed to estimate and describe CPs in different race and sex groups using

**Table 2. Estimates of concurrent relationships by race and sex**

	Total	Men		Women	
		Black	Coloured	Black	Coloured
<b>Six months before survey</b>					
Concurrency among all participants, <i>n</i> (%)	63 (8.4)	25 (13.4)	1 (1.9)	30 (7.8)	7 (5.6)
Concurrency among sexually active participants, <i>n</i> (%)	61 (11.6)	24 (17.8)	1 (3.0)	29 (10.1)	7 (9.7)
<b>Any time in the past year</b>					
Among all participants					
Concurrency, <i>n</i> (%)	138 (18.4)	64 (34.2)	7 (13.0)	58 (15.1)	9 (7.1)
Incidence rate of concurrency, <i>n</i> per 1000 person-years (95% CI)	557.33 (505.18–613.41)	1117.65 (971.25–1279.88)	500.00 (329.50–727.47)	420.37 (357.94–490.55)	166.67 (103.17–254.77)
Among only sexually active participants					
Concurrency, <i>n</i> (%)	134 (25.4)	62 (45.9)	7 (21.2)	56 (19.4)	9 (12.5)
Incidence rate of concurrency, <i>n</i> per 1000 person-years (95% CI)	767.05 (694.15–845.51)	1481.48 (1283.27–1701.64)	818.18 (539.19–1190.41)	545.14 (463.20–637.39)	291.67 (180.55–445.84)
Total number of relationships, med (Range)	1 (1–12)	1 (1–11)	1 (1–12)	1 (1–11)	1 (1–6)
<b>Total number of relationships, <i>n</i> (%)</b>					
1	372 (70.5)	66 (48.9)	23 (69.7)	223 (77.4)	60 (83.3)
2	66 (12.5)	21 (15.6)	2 (6.1)	36 (12.5)	7 (9.7)
≥3	90 (17.0)	48 (35.6)	8 (24.2)	29 (10.1)	5 (6.9)
Number of concurrent relationships, med (Range)	0 (0–11)	0 (0–10)	0 (0–11)	0 (0–10)	0 (0–3)
<b>Number of concurrent relationships, <i>n</i> (%)</b>					
0	394 (74.6)	73 (54.1)	26 (78.8)	232 (80.6)	63 (87.5)
2	72 (13.6)	27 (20.0)	3 (9.1)	36 (12.5)	6 (8.3)
≥3	62 (11.7)	35 (25.9)	4 (12.1)	20 (6.9)	3 (4.2)
Duration of overlap (weeks), med (IQR)	7.5 (2.2–24)	6.75 (2–17)	4 (1–7.6)	9.4 (2–43)	26.5 (16–32)

IQR, inter-quartile range; CI, confidence interval.

a nuanced and holistic definition of CPs. The results of our analysis indicate that not only is there a high incidence and prevalence of CPs in the study communities but the duration of overlapping relationships is also long. Indeed, among all participants we observed a point prevalence of 8.4%, cumulative prevalence of 18.4% and a 7.5-week median duration of overlap among those who had CPs. Importantly, our data also show that most people with two or more relationships in the previous year did not have serially monogamous relationships.

Specifically, among sexually active participants, we see that 45.9% of black men, 19.4% of black women, 21.2% of coloured men and 12.5% of coloured women were engaged in CPs during the past year. This is comparable to the 2009 CP prevalence estimates of Maughan-Brown in his analysis of sexually experienced participants in the Cape Area Panel Study (CAPS): 39% of black men, 14% of black women, 8% of coloured men and 1% of coloured women [30]. Both our study and the CAPS data indicate that black men have the highest frequencies of engaging in CPs. Overall, our study

**Table 3. Binomial regression models for relative risk of concurrency in race–sex groups**

	Model 1 Crude RR (95% CI)	Model 2 Crude RR (95% CI)	Model 3 Adjusted RR (95% CI)	Model 4 Adjusted RR (95% CI)
<b>Sex</b>				
Male	1.00	–	1.00	1.00
Female	0.43 (0.32–0.57)	–	0.43 (0.32–0.57)	0.57 (0.23–1.40)
<b>Race</b>				
Coloured	–	1.00	1.00	1.00
Black	–	1.85 (1.15–2.97)	1.86 (1.17–2.97)	2.16 (1.10–4.27)
Race-sex interaction	–	–	–	0.73 (0.28–1.87)

RR, risk ratio; CI, confidence interval.

**Table 4. Estimates for the effect of sex and race on duration of overlap in concurrent relationships that began in the previous year**

	Model 1 Crude $\beta$ -coefficient (95% CI)	Model 2 Crude $\beta$ -coefficient (95% CI)	Model 3 Adjusted $\beta$ -coefficient (95% CI)	Model 4 Adjusted $\beta$ -coefficient (95% CI)
Intercept	12.99 (8.68–17.31)	14.07 (6.87–21.26)	12.20 (5.70–18.70)	5.89 (3.31–8.47)
Sex				
Female	4.42 (–1.88–10.72)	–	4.43 (–1.87–10.73)	19.43 (7.09–31.76)
Race				
Black	–	0.81 (–7.17–8.78)	0.90 (–6.33–8.13)	8.10 (2.63–13.56)
Race–sex interaction	–	–	–	–17.20 (–31.30–3.10)

CI, confidence interval.

produced higher estimates for all race–sex groups. This may be due to differences in several factors including, but not limited to: the specific areas under study, how the denominators were calculated, how CPs were measured and the mode of interviewing.

Crucially, we see that black men had the highest cumulative prevalence, point prevalence and incidence rates compared to the other three demographics. Black men had 2.16 (95% CI: 1.10–4.27) times the risk of having CPs in the previous year compared to coloured men. More than half of the black men who reported CPs had three or more CPs in the previous year. Kenyon *et al.* report that black people in South Africa have more favourable attitudes towards concurrency than coloured and white people, and this may explain the reasons why prevalence is highest in this group [35]. It has been previously noted that if a specific subgroup of a population is engaging in greater numbers of CPs, such as black men in our population, then the connectedness of the sexual network that they are part of also increases, which may thus increase their HIV transmission probability in a non-linear fashion [31,36].

In addition, all of the race–sex groups investigated in our study have CPs that are characterized by median durations of overlap of four weeks or more. Coloured women had the longest overlaps with the median duration lasting more than six months. This observation is consistent with the finding that coloured women had higher point prevalence but not higher cumulative prevalence of CPs than coloured men. Under the assumption that most coloured women choose coloured men for partners and that it may be easier to hide a concurrent relationship of short duration, the longer overlaps in coloured women versus men may also explain the finding that coloured men were more likely to think that their partners had other partners.

Whether or not to classify the observed durations of overlap as long, is a subjective judgement call. Perhaps the most objective and relevant reference point is the duration of the acute phase of HIV infection. Estimates for the duration of this phase range from one to three months [37,38]. In light of this point of reference, the median durations of overlap observed in our survey are relatively long. The epidemiological implication of long overlaps in relationships is that they may give rise to stable, connected sexual networks because participants go back and forth between sexual partners over

the course of several weeks to months. This means that a person's risk of HIV acquisition will be influenced by the behaviour of others in the network, in addition to his or her own behaviour [39]. These relationships are also more likely to take full advantage of the high viral load and associated high infectiousness during the acute phase of infection, with enhanced transmission potential as a result.

Another key idea that our analysis elucidates is that it is not just men who have high frequencies of engaging in CPs. Large fractions of women also had CPs during the previous year: 19.4% and 12.5% black and coloured sexually active women, respectively. Indeed, out of 65 black women who reported more than one relationship in the previous year, all but nine engaged in concurrency, as opposed to serial monogamy. Many previous studies have reported relatively smaller proportions of women engaging in CPs [30,40,41]. Still, other studies that have investigated the relationship between HIV and CPs have neglected to incorporate CPs from women in their estimates altogether [7,9,10,17]. We argue that in order to understand how or if CPs affects HIV transmission, it is essential to also have accurate estimates of CPs in women and include their contributions to making the sexual network more connected.

Our study has several implications for how research related to estimating the occurrence of CPs is conducted. First, it gives insight into the utility of the UNAIDS proposed indicator of measuring the point prevalence of CPs six months before the survey. As expected, based on the relationship overlap data, we found large differences between the reported point prevalence and cumulative prevalence, especially for men. Men in our study had shorter durations of overlap and so cross-sectional snapshots of the sexual network (i.e. point prevalence estimates) were less likely to include the short concurrent relationships. This implies that the point prevalence alone may not be a sufficient proxy for studies investigating the relationship between HIV and CPs. Linking the cumulative prevalence indicator to the point prevalence indicator and duration of overlap can help to distinguish between different types of CP patterns at the population level. For example, if point prevalence and cumulative prevalence are small, and the average duration of overlap is short, relative to the period of observation, this would suggest that engaging in CPs is confined to a small subgroup of individuals. Large differences between point prevalence and cumulative

prevalence, despite relatively short mean duration of overlap, suggest that there is a large fraction of people infrequently engaging in CPs.

Measuring the incidence of CPs may add value to studies that estimate the effect of behavioural change interventions. For instance, there may be settings in which a small core group of individuals is engaging in a high number of short CPs each year. If the intervention reduces the rate of acquiring CPs, but not the size of this core group, then the one-year cumulative prevalence of CPs before and after the intervention would be similar. Moreover, the short duration of overlap would result in point prevalence estimates of CPs that are very small, and therefore may lack statistical power to detect differences over time. Our study suggests that point prevalence, incidence, cumulative prevalence, number of relationships and duration of overlap, in combination, tell a more complete story. High proportions of people engaging in CPs (i.e. cumulative prevalence), high frequency of engaging in CPs (i.e. incidence) and long overlaps would lead to a highly connected network and faster spread of HIV.

Second, our study offers a better way to measure concurrency for future studies that investigate the relationship between CPs and HIV [42]. Although UNAIDS suggests an indicator that researchers should use, it does not elaborate on how the source data should be collected. The largest strength of our study is that our estimates of CPs were derived from a questionnaire that used ACASI, which has been shown to increase reporting of sensitive sexual behaviours in other African contexts [43–46] and reduce social desirability bias in our own study [25]. Furthermore, the questionnaire was presented on a touchscreen desktop computer, which exhibited a visual timeline with dates during the previous year. The durations for each relationship were constantly displayed below the timeline in different colours to help participants recall new relationships in the context of dates for previous relationships they enumerated. Visual timelines have been shown to foster internal consistency in reporting relationship dates [47,48]. Moreover, we believe our questionnaire reduced fatigue bias because it was designed specifically to ask relatively few questions about sexual behaviours in each relationship. This contrasts with DHSs, for example, which have been previously used to estimate CP prevalence in different populations [2,49]. Also, we believe our more accurate CPs estimates lend themselves well as calibration data for modelling studies of the spread of HIV and effectiveness of combination prevention strategies in an urban South African context.

Our study is not without limitations. First, our CPs estimates may still be biased despite using an ACASI questionnaire. A study conducted on Likoma Island in Lake Malawi, using socio-centric data to look at inter-partner agreement of reports of sexual relationships, indicated that CPs estimates may still be unreliable even when ACASI questionnaires are administered [50]. Furthermore, the same study also found that relationships of longer duration were more likely to be reported and short-term relationships may be underreported by participants. Date heaping is another form of bias that may be present, resulting in overestimates of concurrency [51]. However, we believe our study was less likely to show these

forms of bias due to the visual timeline that we used to help people accurately recall their relationships and their durations. A second limitation of our study is that our data were left truncated, and therefore we do not have the beginning dates of relationships that started one year before the survey. This means that the median duration of overlap we calculated in Table 2 is probably an underestimate of the true median, because it is possible that two relationships that were overlapping at the start of the previous 12-month window, had been overlapping before the period of observation. The fact that the data were left truncated limited the number of relationships we could use for our analysis of the effect of sex and race on duration of overlap in CPs. It also prevented us from performing more advanced time-to-event analyses for the risk of entering CPs.

Although we do not have any evidence to advocate for or against CPs interventions, we believe that it is premature to give up on behavioural interventions as a mode for HIV prevention. The lack of evidence for the effectiveness of CPs-reduction interventions does not necessarily mean that such interventions are ineffective. Many reasons for this lack of evidence exist. In the current context of a strong international focus on biomedical interventions for HIV prevention, behavioural research is relatively underfunded [52]. As a result, studies with a primary focus on sexual behaviour lack funds to collect HIV biomarkers [53]. Alternatively, many studies with a primary focus on biomedical interventions include the measurement of HIV incidence but studies of sexual relationship histories are often treated as secondary analyses [54], and suboptimal designs for collecting this sensitive information are used.

## Conclusions

Although we cannot provide evidence of CPs influencing HIV transmission, we do offer a useful way forward for measuring and defining CPs for future studies. We have demonstrated that in economically disadvantaged areas around Cape Town, CPs rates are high and the average duration of overlaps are relatively long. The conditions needed to create and maintain a highly connected sexual network have been met in this setting. Our study does not lend itself to providing evidence for the commencement, continuation or termination of public health interventions related to CPs. However, our estimates may be useful in future modelling studies that attempt to improve our understanding of what combination of CPs incidence, prevalence and duration of overlap would be sufficient to result in sizeable increases in the rate of HIV transmission. We believe it would be useful to repeat our survey and analysis in other settings with varying degrees of HIV prevalence to see if further associations can be found with our proposed suite of CP indicators and HIV.

### Authors' affiliations

<sup>1</sup>The South African Department of Science and Technology/National Research Foundation (DST/NRF) Centre of Excellence in Epidemiological Modelling and Analysis (SACMA), Stellenbosch University, Stellenbosch, South Africa; <sup>2</sup>International Centre for Reproductive Health, Ghent University, Ghent, Belgium; <sup>3</sup>Center for Statistics, Hasselt University, Diepenbeek, Belgium; <sup>4</sup>Centre for Health Economic Research and Modeling Infectious Diseases, Vaccine and Infectious Disease Institute, University of Antwerp, Wilrijk, Belgium

#### Competing interests

There are no competing interests.

#### Authors' contributions

RB, WD and NH jointly designed the ACASI survey. RB coordinated data collection. RB conducted the statistical analysis. NH and WD supervised the data analysis. WD and RB wrote the first draft manuscript. All authors contributed during the editing process and approved the final, submitted manuscript.

#### Acknowledgements

The authors thank the survey respondents for their participation in this study. They are also indebted to the research assistants for countless hours of dedicated fieldwork and to the Desmond Tutu TB Centre at Stellenbosch University for their guidance in coordination and planning of this cross-sectional sexual behaviour survey. The Flemish Interuniversity Council and the Research Foundation – Flanders (FWO) provided financial support for the survey. WD and RB received funding from the Canadian International Development Agency and FWO. NH is grateful for the support from the University of Antwerp scientific chair in Evidence-Based Vaccinology, financed from 2009 to 2014 by an unrestricted donation from Pfizer.

#### References

1. Eaton JW, Hallett TB, Garnett GP. Concurrent sexual partnerships and primary HIV infection: a critical interaction. *AIDS Behav.* 2011;15(4):687–92.
2. Fox AM. Marital concurrency and HIV risk in 16 African countries. *AIDS Behav.* 2014;18(4):791–800.
3. Goodreau SM, Cassels S, Kasprzyk D, Montano DE, Greek A, Morris M. Concurrent partnerships, acute infection and HIV epidemic dynamics among young adults in Zimbabwe. *AIDS Behav.* 2012;16(2):312–22.
4. Kenyon C, Colebunders R. Strong association between point-concurrency and national peak HIV prevalence. *Int J Infect Dis.* 2012;16(11):e826–7.
5. Kretzschmar M, Carael M. Is concurrency driving HIV transmission in Sub-Saharan African sexual networks? The significance of sexual partnership typology. *AIDS Behav.* 2012;16:1746–52.
6. UNAIDS Reference Group on Estimates Modelling and Projections. Working group on measuring concurrent sexual partnerships. HIV: consensus indicators are needed for concurrency. *Lancet.* 2010;375(9715):621–2.
7. Kenyon C, Buyze J. Correlation between national peak HIV prevalence and concurrency, STI treatment capability, condom use and circumcision: an ecological study. *World J AIDS.* 2014;4:249–57.
8. Kenyon CR, Osbak K, Buyze J. The prevalence of HIV by ethnic group is correlated with HSV-2 and Syphilis prevalence in Kenya, South Africa, the United Kingdom, and the United States. *Interdiscip Perspect Infect Dis.* 2014;2014:284317.
9. Tanser F, Barnighausen T, Hund L, Garnett GP, McGrath N, Newell ML. Effect of concurrent sexual partnerships on rate of new HIV infections in a high-prevalence, rural South African population: a cohort study. *Lancet.* 2011;378(9787):247–55.
10. Westercamp N, Mattson CL, Bailey RC. Measuring prevalence and correlates of concurrent sexual partnerships among young sexually active men in Kisumu, Kenya. *AIDS Behav.* 2013;17(9):3124–32.
11. Reniers G, Watkins S. Polygyny and the spread of HIV in sub-Saharan Africa: a case of benign concurrency. *AIDS.* 2010;24(2):299–307.
12. Sowers L, Isaac AG, Stillwaggon E. HIV and concurrent sexual partnerships: modelling the role of coital dilution. *J Int AIDS Soc.* 2011;14:44.
13. Gaydos L, Reniers G, Hellinginger S. Partnership concurrency and coital frequency. *AIDS Behav.* 2013;17(7):2376–86.
14. Delva W, Meng F, Beaulair R, Deprez N, Temmerman M, Welte A, et al. Coital frequency and condom use in monogamous and concurrent sexual relationships in Cape Town, South Africa. *J Int AIDS Soc.* 2013;16:18034. doi:<http://dx.doi.org/10.7448/IAS.16.1.18034>
15. Go MH, Blower S. What impact will reducing concurrency have on decreasing the incidence of HIV in heterosexual populations? *Sex Transm Dis.* 2012;39(6):414–5.
16. Chopra M, Townsend L, Johnston L, Mathews C, Tomlinson M, O'Bra H, et al. Estimating HIV prevalence and risk behaviors among high-risk heterosexual men with multiple sex partners: use of respondent-driven sampling. *J Acquir Immune Defic Syndr.* 2009;51(1):72–7.
17. Sully E, Reniers G, Ekoru K, Seeley J, editors. Partnership concurrency and HIV incidence in a population-based cohort study in rural Uganda. San Francisco, CA: Population Association of America 2012 Annual Meeting; 2012.
18. Maughan-Brown B. Variation in concurrent sexual partnerships and sexually transmitted diseases among african men in cape town, South Africa. *Sex Transm Dis.* 2012;39(7):537–42.
19. Hellinginger S, Mkwandawire J, Kohler HP. A new approach to measuring partnership concurrency and its association with HIV risk in couples. *AIDS Behav.* 2014;18(12):2291–301.
20. Eaton JW, Takavarasha FR, Schumacher CM, Mugurungi O, Garnett GP, Nyamukapa C, et al. Trends in concurrency, polygyny, and multiple sex partnerships during a decade of declining HIV prevalence in Eastern Zimbabwe. *J Infect Dis.* 2014;210(Suppl 2):S562–8.
21. Kenyon C, Buyze J, Colebunders R. HIV prevalence by race co-varies closely with concurrency and number of sex partners in South Africa. *PLoS One.* 2013;8(5):e64080.
22. Adimora AA, Hughes JP, Wang J, Haley DF, Golin CE, Magnus M, et al. Characteristics of multiple and concurrent partnerships among women at high risk for HIV infection. *J Acquir Immune Defic Syndr.* 2014;65(1):99–106.
23. Morris M, Epstein H, Wawer M. Timing is everything: international variations in historical sexual partnership concurrency and HIV prevalence. *PLoS One.* 2010;5(11):e14092.
24. Kwna ZA, Mwanzo IJ, Bukusi EA, Achiro LF, Shisanya CA. A cross-sectional survey of prevalence and correlates of couple sexual concurrency among married couples in fishing communities along Lake Victoria in Kisumu, Kenya. *Sex Transm Infect.* 2014;90(2):139–44.
25. Beaulair R, Meng F, Deprez N, Temmerman M, Welte A, Hens N, et al. Evaluating audio computer assisted self-interviews in urban South African communities: evidence for good suitability and reduced social desirability bias of a cross-sectional survey on sexual behaviour. *BMC Med Res Methodol.* 2013;13:11.
26. Lurie MN, Rosenthal S. Concurrent partnerships as a driver of the HIV epidemic in Sub-Saharan Africa? The evidence is limited. *AIDS Behav.* 2010;14:17–24.
27. Shisanya O. HIV/AIDS in South Africa: at last the glass is half full. 6th South African AIDS Conference; 2013 June 18–21; Durban, South Africa.
28. Myer L, Ehrlich RI, Susser ES. Social epidemiology in South Africa. *Epidemiol Rev.* 2004;26:112–23.
29. Mah TL. Prevalence and correlates of concurrent sexual partnerships among young people in South Africa. *Sex Transm Dis.* 2010;37(2):105–8.
30. Maughan-Brown B. Concurrent sexual partnerships among young adults in Cape Town, South Africa: how is concurrency changing? *Sex Health.* 2013;10(3):246–52.
31. Kenyon C, Dlamini S, Boule A, White R, Badr M. A network-level explanation for the differences in HIV prevalence in South Africa's racial groups. *Afr J AIDS Res.* 2009;8(3):243–54.
32. City of Cape Town. City of Cape Town-census 2001. Cape Town: Strategic Development Information; 2001.
33. Ayles HM, Sismanidis C, Beyers N, Hayes RJ, Godfrey-Faussett P. ZAMSTAR, The Zambia South Africa TB and HIV reduction study: design of a 2 × 2 factorial community randomized trial. *Trials.* 2008;9:63.
34. Delva W, Beaulair R, Welte A, Vansteelandt S, Hens N, Aerts M, et al. Age-disparity, sexual networkedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol. *BMC Public Health.* 2011;11:616.
35. Kenyon CR, Osbak K, Buyze J, Johnson S, van Lankveld J. Variations of sexual scripts relating to concurrency by race, class, and gender in South Africa. *J Sex Res.* 2014:1–9. [Epub ahead of print].
36. Morris M, Kretzschmar M. Concurrent partnerships and the spread of HIV. *AIDS.* 1997;11(5):641–8.
37. Cohen MS, Gay CL, Busch MP, Hecht FM. The detection of acute HIV infection. *J Infect Dis.* 2010;202(Suppl 2):S270–7.
38. Hollingsworth TD, Anderson RM, Fraser C. HIV-1 transmission, by stage of infection. *J Infect Dis.* 2008;198(5):687–93.
39. Epstein H, Morris M. Concurrent partnerships and HIV: an inconvenient truth. *J Int AIDS Soc.* 2011;14:13.
40. Maher D, Waswa L, Karabarinde A, Baisley K. Concurrent sexual partnerships and associated factors: a cross-sectional population-based survey in a rural community in Africa with a generalised HIV epidemic. *BMC Public Health.* 2011;11:651.
41. Xu H, Luke N, Zulu EM. Concurrent sexual partnerships among youth in urban Kenya: prevalence and partnership effects. *Popul Stud.* 2010;64(3):247–61.

42. Mah TL, Shelton JD. Concurrency revisited: increasing and compelling epidemiological evidence. *J Int AIDS Soc.* 2011;14:33.
43. Dolezal C, Marhefka SL, Santamaria EK, Leu CS, Brackis-Cott E, Mellins CA. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Archiv Sex Behav.* 2012;41(2):401–10.
44. Hewett P, Mensch B, Erukar A. Consistency in the reporting of sexual behavior by adolescent girls in Kenya: a comparison of interviewing methods. *Sex Transm Infect.* 2004;80(Suppl 2):iii43–8.
45. Kissinger P, Rice J, Farley T, Trim S, Jewitt K, Margavio V, et al. Application of computer-assisted interviews to sexual behavior research. *Am J Epidemiol.* 1999;149(10):950–4.
46. Phillips AE, Gomez GB, Bolly MC, Garnett GP. A systematic review and meta-analysis of quantitative interviewing tools to investigate self-reported HIV and STI associated behaviours in low- and middle-income countries. *Int J Epidemiol.* 2010;39(6):1541–55.
47. Kabiru CW, Luke N, Itugbara CO, Zulu E. The correlates of HIV testing and impacts on sexual behaviour: evidence from a life history study of young people in Kisumu, Kenya. *BMC Public Health.* 2010;10:412.
48. Luke N, Clark S, Zulu E. The relationship history calendar: improving sexual behaviour data among youth in developing country settings. Population Association of America (PAA) Conference; 2008 April 17–19; New Orleans, USA. 2008.
49. Reniers G, T'faily R. Polygyny, partnership concurrency, and HIV transmission in Sub-Saharan Africa. *Demography.* 2012;49:1075–101.
50. HELLERINGER S, KOHLER HP, KALLIANI-PHIRI L, MKANDAWIRE J, ARMBRUSTER B. The reliability of sexual partnership histories: implications for the measurement of partnership concurrency during surveys. *AIDS.* 2011;25(4):503–11.
51. Morris M, O'Gorman J. The impact of measurement error on survey estimates of concurrent partnerships. *Math Pop Stud.* 2000;8(3):231–49.
52. Schwartlander B, Stover J, Hallett T, Atun R, Avila C, Gouws E, et al. Towards an improved investment approach for an effective response to HIV/AIDS. *Lancet.* 2011;377(9782):2031–41.
53. Coates TJ, Richter L, Caceres C. Behavioural strategies to reduce HIV transmission: how to make them work better. *Lancet.* 2008;372(9639):669–84.
54. MacQueen KM. Framing the social in biomedical HIV prevention trials: a 20-year retrospective. *J Int AIDS Soc.* 2011;14(Suppl 2):S3.

# APPENDIX E

## why?



The reason for this study was to understand and record the **difference in age** in male/female relationships



This study also wanted to find out more about **relationships** and why **women choose certain men** for relationships.



The study wanted to find out if being in **two or more relationships** ("concurrent relationships") at the same time could make it **easier for HIV to be transmitted**.

## when?

between **July 2011** and **February 2012**

## perceived risks of dating older men and age-mates

### most women thought:

- Older men are not risky partners
- Men who were the same age or younger were not ideal, because they could be disruptive and abusive

### a few women thought:

- Older men were less likely to use condoms
- Older men were more likely to cheat on them

Women wanted to **avoid abuse and disrespect** and the often prompted them to **send a particular relationship** and start a different, new relationship instead.

### perceived risks of dating an older man

"The more you think about it, don't think there is one who is not a risk." (Age 23, Black)

"No, like, there is no risk with being with an older man." (Age 47, colored)

"They can't send a swab for HIV." (Age 23, Black)

"This [young girl] is not a risk because she is trying to get as long as she can with him, and she can be involved with many other children." (Age 51, Black)

### perceived risks of dating a same age man

"The young ones, they like being abusive and disrespectful." (Age 44, colored)

"The younger guys, they're disrespectful, they can't help to very much." (Age 26, Black)

"If you're dating a younger guy, you're gonna get a lot of reasons why you're not a risk because it's because of the type of guy you're dating." (Age 26, Black)

"You're not really just looking for a boy, because you're looking for a girl that you can have sex with." (Age 21, Black)

### reasons for relationship gap

"It's a very young one and he would give me a lot of reasons why he's not like that." (Age 59, colored)

"I want to normally just have sex with a boy, but because my ex-boyfriend had sex and had a girl, I had to send him." (Age 26, colored)

"Maybe the person, he becomes distant if you don't have sex with him." (Age 26, colored)

"I want to have sex with a boy, but because my ex-boyfriend had sex and had a girl, I had to send him." (Age 21, Black)

## HIV prevention programming and future research

- The unmet research questions for questions worked really well for participants who were answering
- The unmet research questions that needs to be done in this area to understand the deeper root causes of risky sexual behaviors.
- The unmet research questions need to be done in this area to understand the deeper root causes of HIV risk behaviors.
- The unmet research questions need to be done in this area to understand the deeper root causes of HIV risk behaviors.



## Who?

**878**

Participants took the survey

**750** heterosexual between 15 & 70

years of age

**average age:** 35 years of age

**68%** were **women**

**76%** were **black** and **24%** were **coloured**

**79%** were **unemployed**

**59%** had at least a **secondary education**

## Where?

**46%**

Whitehead

**29%**

Dalhousie

**25%**

Kloofbaai Site C

(age limit, social class)



## How?

People from the **Zambia South Africa YS and AIDS Reduction study (ZAMSTAR)** were asked if they could answer some questions.

**23** women also took the Personal Health Survey



## Types of Questions

We collected information about:

- Personal information (age, education, etc.)
- Relationships
- Partner age
- Whether they thought their partner had sex with others at the same time as them
- Whether or not the participant or their partner had sex with others at the same time as their first intercourse.
- How far away the partner lived
- How many times they had sex
- What they thought was good sex
- Conditions
- Condom use
- Total number of sexual partners they ever had
- When they started having sex
- When they started when answering questions
- What they thought was good sex
- Conditions
- Condom use
- Total number of sexual partners they ever had

**878** participants took the survey and reported **1,128** relationships



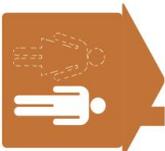
The heterosexual black and coloured partners between the ages of **15 and 70** reported **1,003** relationships

- Only **51%** of those relationships involved sexual intercourse
- 44%** started in the year before the participant took the survey

Among relationships that started in the past year,

**82%** did not live with their sexual partner

**33%** lived in a different township or province



**25%** of the sexual active participants, reported sex with all partners in the past year

**29%** of these recent relationships, the participant was under 25 years of age

**86%** of people who were reported to have sex with their partner, reported overlapping relationships



**27%** of these relationships overlapped with another partner or partner in the last 12 months

**25%** of sexually active participants had one relationship 'another' in the past year

| on average, relationships overlapped for 1.5 weeks |
|--|--|--|--|--|--|
| ✓  | ✓  | ✓  | ✓  | ✓  | ✓  |
| ✓  | ✓  | ✓  | ✓  | ✓  | ✓  |
| ✓  | ✓  | ✓  | ✓  | ✓  | ✓  |
| ✓  | ✓  | ✓  | ✓  | ✓  | ✓  |
| ✓  | ✓  | ✓  | ✓  | ✓  | ✓  |
| ✓  | ✓  | ✓  | ✓  | ✓  | ✓  |

**16%** of all relationships reported with partners who were **5-9 years younger** with partners **15%** were **10 or more younger**

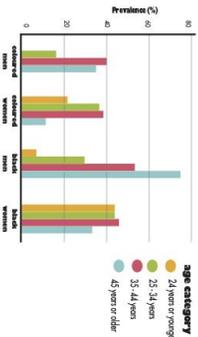
**33%** of all relationships reported with partners who were **5-9 years older** with partners **16%** were **10 or more older**

For **39%** of relationships reported, participants thought their partners had another partner. **76%** of partners were reported to "share partners" and **39%** were "casual partners"

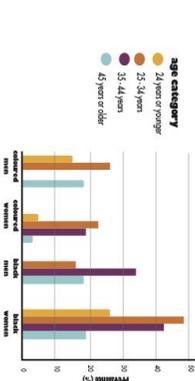
## Risky Behaviour:

- Those under the influence of alcohol at the start of a relationship have **2 times** the risk of inconsistent condom use
- Women were **2.3 times** less likely to have a concurrent relationship than men
- Black participants were **1.9 times** more likely than coloured participants to have a concurrent relationship
- Those who had a higher sex frequency had greater risk of beginning a concurrent relationship
- Participants to have a concurrent relationship
- Black participants were **2.9 times** more likely than coloured participants to have a concurrent relationship
- Those who had a higher sex frequency had greater risk of beginning a concurrent relationship

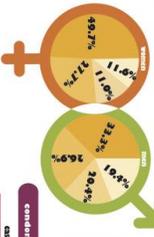
## Age-disparate relationships



## Prevalence of HIV



## Prevalence of overlapping relationship types



types of overlapping relationships

- main partner and overlap is greater than 1 month
- 1 month overlap or less
- casual partner and overlap is greater than 1 month
- occasional

## condom use in different partner types



## REFERENCES

---

1. Unaid. Global AIDS Update 2016. 2016;
2. UNAIDS. AIDSinfo [Internet]. 2017 [cited 2017 Jun 14]. Available from: <http://www.unaids.org/en/dataanalysis/datatools/aidsinfo/>
3. Abdool-Karim Q, Abdool-Karim SS. The evolving HIV epidemic in South Africa. *Int J Epidemiol.* 2002/03/27. 2002;31(1):37–40.
4. Abdool Karim S, Churchyard G, Abdool Karim Q, Lawn S. HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. *Lancet.* 2009;374:921–33.
5. UNAIDS. Country: Malawi [Internet]. UNAIDS Country factsheets. 2018. Available from: <http://www.unaids.org/en/regionscountries/countries/malawi>
6. De Cock KM, Crowley SP, Lo YR, Granich RM, Williams BG. Preventing HIV transmission with antiretrovirals. *Bull World Heal Organ.* 2009/08/04. 2009;87(7):488–488A.
7. Granich R, Crowley S, Vitoria M, Lo YR, Souteyrand Y, Dye C, et al. Highly active antiretroviral treatment for the prevention of HIV transmission. *J Int AIDS Soc.* 2010;13:1.
8. Delva W, Abdool Karim Q. The HIV epidemic in Southern Africa - Is an AIDS-free generation possible? *Curr HIV/AIDS Rep.* 2014;11(2):99–108.
9. Dellar RC, Dlamini S, Karim QA. Adolescent girls and young women: Key populations for HIV epidemic control. *J Int AIDS Soc.* 2015;18(2):64–70.
10. Anderson RM, May RM, Ng TW, Rowley JT. Age-Dependent Choice of Sexual Partners and the Transmission Dynamics of HIV in Sub-Saharan Africa. *Philos Trans Biol Sci.* 1992;336(1277):135–55.
11. Chapman R, White RG, Shafer LA, Pettifor A, Mugurungi O, Ross D, et al. Do behavioural differences help to explain variations in HIV prevalence in adolescents in sub-Saharan Africa? *Trop Med Int Heal.* 2010;15(5):554–66.
12. Garnett GP, Anderson RM. Sexually transmitted diseases and sexual behavior: insights from mathematical models. *J Infect Dis.* 1996;174 Suppl:S150-61.
13. Hallett TB, Gregson S, Lewis JJ, Lopman BA, Garnett GP. Behaviour change in generalised HIV epidemics: impact of reducing cross-generational sex and delaying age at sexual debut. *Sex Transm Infect.* 2007/02/23. 2007;83 Suppl 1:i50-54.
14. Leclerc-Madlala S. Age-disparate and intergenerational sex in southern Africa: the dynamics of hypervulnerability. *AIDS.* 2008/12/17. 2008;22 Suppl 4:S17-25.
15. Newman MEJ. Assortative Mixing in Networks. *Phys Rev Lett.* 2002;89.
16. Viera AJ. Odds ratios and risk ratios: what's the difference and why does it matter? *South Med J.* 2008 Jul;101(7):730–4.

17. Davies HT, Crombie IK, Tavakoli M. When can odds ratios mislead? *BMJ*. 1998 Mar 28;316(7136):989–91.
18. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
19. Shisana O, Rehle T, Simbayi L, Zuma K, Jooste S, Zungu N, et al. *South African National HIV Prevalence, Incidence and Behaviour Survey, 2012*. Cape Town; 2014.
20. Office NS, Macro ICF. *Malawi Demographic and Health Survey 2010*. Zomba, Malawi Calverton, Maryland, USA: NSO and ICF Macro; 2011.
21. Helleringer S, Kohler H-P, Mkandawire J. Women underestimate the age of their partners during survey interviews: implications for HIV risk associated with age mixing in northern Malawi. *Sex Transm Dis*. 2011;38(11):1030–5.
22. Gouws E, Stanecki KA, Lyster R, Ghys PD. The epidemiology of HIV infection among young people aged 15-24 years in southern Africa. *Aids*. 2008;22(SUPPL. 4):S5–16.
23. Pettifor AE, Rees H V, Kleinschmidt I, Steffenson AE, MacPhail C, Hlongwa-Madikizela L, et al. Young people's sexual health in South Africa: HIV prevalence and sexual behaviors from a nationally representative household survey. *AIDS*. 2005/09/02. 2005;19(14):1525–34.
24. Delva W, Meng F, Tolentino S, Hens N, Welte A, Dushoff J. Age mixing and sustained HIV epidemics: not the size but the variation of age gaps counts. In: *Epidemics4 – 4th International Conference on Infectious Diseases Dynamics*, 19-22 November 2013. Amsterdam, The Netherlands; 2013.
25. Longfield K, Glick A, Waithaka M, Berman J. Relationships between older men and younger women: Implications for STIs/HIV in Kenya. *Stud Fam Plann*. 2004;35(2).
26. Luke N. Confronting the 'sugar daddy' stereotype: age and economic asymmetries and risky sexual behavior in urban Kenya. *Int Fam Plan Perspect*. 2005/05/13. 2005;31(1):6–14.
27. Wingood GM, Diclemente RJ. Application of the Theory of Gender and Power to Examine HIV-Related Exposures , Risk Factors , and Effective Interventions for Women. 2000;27(October):539–65.
28. Woolf SE, Maisto SA. Gender differences in condom use behavior? The role of power and partner-type. *Sex Roles*. 2008;58(9–10):689–701.
29. Morrison-Beedy D, Xia Y, Passmore D. Sexual risk factors for partner age discordance in adolescent girls and their male partners. *J Clin Nurs*. 2013;22(23–24):3289–99.
30. Nkosana J, Nkosana LBM. Intergenerational sexual relationships: The voices of men. *Integr J Br*. 2015;2(9):14–31.
31. 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*. 2002/06/12. 2002;359(9321):1896–903.

32. Maughan-brown B, Evans M, George G. Sexual Behaviour of Men and Women within Age-Disparate Partnerships in South Africa : Implications for Young Women ' s HIV Risk. *PLoS One*. 2016;1–16.
33. Ritchwood TD, Hughes JP, Jennings L, MacPhail C, Williamson B, Selin A, et al. Characteristics of Age-Discordant Partnerships Associated With HIV Risk Among Young South African Women (HPTN 068). *JAIDS J Acquir Immune Defic Syndr*. 2016 Aug;72(4):423–9.
34. Toska E, Cluver LD, Boyes M, Pantelic M, Kuo C. From 'sugar daddies' to 'sugar babies': exploring a pathway between age-disparate sexual relationships, condom use, and adolescent pregnancy in South Africa. *Sex Health*. 2016;12(1):59–66.
35. Sturdevant MS, Belzer M, Weissman G, Friedman LB, Sarr M, Muenz LR. The relationship of unsafe sexual behavior and the characteristics of sexual partners of HIV infected and HIV uninfected adolescent females. *J Adolesc Heal*. 2001;29(3 SUPPL. 1):64–71.
36. Wyrod R, Fritz K, Woelk G, Jain S, Kellogg T, Chirowodza A, et al. Beyond sugar daddies: Intergenerational sex and AIDS in Urban Zimbabwe. *AIDS Behav*. 2011;15(6):1275–82.
37. Maughan-Brown B, Kenyon C, Lurie MN. Partner Age Differences and Concurrency in South Africa: Implications for HIV-Infection Risk Among Young Women. *AIDS Behav*. 2014;
38. Street RA, Reddy T, Ramjee G. The generational effect on age disparate partnerships and the risk for human immunodeficiency virus and sexually transmitted infections acquisition. *Int J STD AIDS*. 2015;
39. UNAIDS Reference Group on Estimates Modelling and Projections. Working Group on Measuring Concurrent Sexual Partnerships. HIV: consensus indicators are needed for concurrency. *Lancet*. 2009/12/04. 2010;375(9715):621–2.
40. Beauclair R, Hens N, Delva W. Concurrent partnerships in Cape Town, South Africa: race and sex differences in prevalence and duration of overlap. *J Int AIDS Soc*. 2015;18(1):19372.
41. 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(15):2245–52.
42. Morris M, Epstein H, Wawer M. Timing is Everything: International Variations in Historical Sexual Partnership Concurrency and HIV Prevalence. *PLoS One*. 2010;5(11):e14092.
43. Sandoy IF, Dzekedzeke K, Fylkesnes K. Prevalence and correlates of concurrent sexual partnerships in Zambia. *AIDS Behav*. 2008/10/09. 2008;14(1):59–71.
44. Boily MC, Baggaley RF, Wang L, Masse B, White RG, Hayes RJ, et al. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *Lancet Infect Dis*. 2009/01/31. 2009;9(2):118–29.
45. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, Laeyendecker O, et al. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai,

- Uganda. *J Infect Dis.* 2005/04/06. 2005;191(9):1403–9.
46. Eaton JW, Hallett TB, Garnett GP. Concurrent sexual partnerships and primary HIV infection: A critical interaction. *AIDS Behav.* 2011;15(4):687–92.
  47. Fox AM. Marital Concurrence and HIV Risk in 16 African Countries. *AIDS Behav.* 2014/01/09. 2014;18(4):791–800.
  48. Goodreau SM, Cassels S, Kasprzyk D, Montano DE, Greek A, Morris M. Concurrent partnerships, acute infection and HIV epidemic dynamics among young adults in Zimbabwe. *AIDS Behav.* 2012;16(2):312–22.
  49. Kenyon C, Colebunders R. Strong association between point-concurrency and national peak HIV prevalence. *Int J Infect Dis.* 2012;16(11):e826-7.
  50. Kretzschmar M, Caraël M. Is concurrency driving HIV transmission in sub-Saharan african sexual networks? the significance of sexual partnership typology. *AIDS Behav.* 2012;16(7):1746–52.
  51. Kenyon C, Buyze J. Correlation between National Peak HIV Prevalence and Concurrency, STI Treatment Capability, Condom Use and Circumcision: An Ecological Study. *World J AIDS.* 2014;4:249–57.
  52. Kenyon CR, Osbak K, Buyze J. The Prevalence of HIV by Ethnic Group Is Correlated with HSV-2 and Syphilis Prevalence in Kenya, South Africa, the United Kingdom, and the United States. *Interdiscip Perspect Infect Dis.* 2014;2014:284317.
  53. Tanser F, Barnighausen T, Hund L, Garnett GP, McGrath N, Newell ML. Effect of concurrent sexual partnerships on rate of new HIV infections in a high-prevalence, rural South African population: a cohort study. *Lancet.* 2011/07/19. 2011;378(9787):247–55.
  54. Westercamp N, Mattson CL, Bailey RC. Measuring prevalence and correlates of concurrent sexual partnerships among young sexually active men in Kisumu, Kenya. *AIDS Behav.* 2013/03/28. 2013;17(9):3124–32.
  55. Wawer MJ, Sewankambo NK, Berkley S, Serwadda D, Musgrave SD, Gray HR, et al. Incidence of HIV-1 infection in a rural region of Uganda. *BMJ.* 1994;308(6922).
  56. Hyman JM, Stanley EA. Using mathematical models to understand the AIDS epidemic. *Math Biosci.* 1988;90(1–2):415–73.
  57. Todd J, Grosskurth H, Changalucha J, Obasi A, Mosha F, Balira R, et al. Risk factors influencing HIV infection incidence in a rural African population: a nested case-control study. *J Infect Dis.* 2006;193(3):458–66.
  58. Weiser SD, Leiter K, Heisler M, McFarland W, Percy-De Korte F, DeMonner SM, et al. A population-based study on alcohol and high-risk sexual behaviors in Botswana. *PLoS Med.* 2006;3(10):1940–8.
  59. Chersich MF, Rees H V, Scorgie F, Martin G. Enhancing global control of alcohol to reduce unsafe sex and HIV in sub-Saharan Africa. *Glob Heal.* 2009/11/19. 2009;5:16.
  60. Fisher JC, Bang H, Kapiga SH. The association between HIV infection and

- alcohol use: a systematic review and meta-analysis of African studies. *Sex Transm Dis.* 2007/12/01. 2007;34(11):856–63.
61. Pithey A, Parry C. between alcohol use and HIV infection Alcohol use and HIV infection , disease progression and treatment Alcohol use and HIV sexual risk behaviours. *Methodology.* 2009;6(4):155–69.
  62. Morojele NK, Kachieng'a MA, Mokoko E, Nkoko MA, Parry CDH, Nkowane AM, et al. Alcohol use and sexual behaviour among risky drinkers and bar and shebeen patrons in Gauteng province, South Africa. *Soc Sci Med.* 2006;62(1):217–27.
  63. Peltzer K, Ramlagan S. Alcohol use trends in South Africa. *J Soc Sci.* 2009;18(1):1–12.
  64. E. Greig F, Koopman C. Multilevel Analysis of Women's Empowerment and HIV Prevention: Quantitative Survey Results from a Preliminary Study in Botswana. *AIDS Behav.* 2003;7(2).
  65. Kalichman SC, Simbayi LC, Kaufman M, Cain D, Jooste S. Alcohol use and sexual risks for HIV/AIDS in sub-saharan Africa: Systematic review of empirical findings. *Prev Sci.* 2007;8(2):141–51.
  66. Kiene S, Simbayi LC, Abrams A. High rates of unprotected sex occurring among HIV-positive individuals in a daily diary study in South Africa: the role of alcohol use. *J Acquir Immune Defic Syndr.* 2008;49(2):219–26.
  67. Kalichman SC, Simbayi LC, Cain D, Jooste S. Heterosexual anal intercourse among community and clinical settings in Cape Town, South Africa. *Sex Transm Infect.* 2009/05/12. 2009;85(6):411–5.
  68. Chersich MF, Rees H V. Vulnerability of women in southern Africa to infection with HIV: biological determinants and priority health sector interventions. *AIDS.* 2008/12/17. 2008;22 Suppl 4:S27-40.
  69. Karim SA, Ramjee G. Anal Sex and HIV Transmission in Women. *Am J Public Health.* 1998;88(8):1265–6.
  70. Varghese B, Maher JE, Peterman TA, Branson BM, Steketee RW. Reducing the risk of sexual HIV transmission: quantifying the per-act risk for HIV on the basis of choice of partner, sex act, and condom use. *Sex Transm Dis.* 2002;29(1):38–43.
  71. Ndinda C, Chimbwete C, McGrath N, Pool R. Perceptions of anal sex in rural South Africa. *Cult Health Sex.* 2008;10(2):205–12.
  72. Yi TJ, Shannon B, Prodder J, McKinnon L, Kaul R. Genital immunology and HIV susceptibility in young women. *Am J Reprod Immunol.* 2013;69 Suppl 1:74–9.
  73. Rakhi J, Sumathi M. Contraceptive Methods: Needs, Options and Utilization. 2012;
  74. Stanley M. Early age of sexual debut: a risky experience. *J Fam Plann Reprod Heal Care.* 2009;35(2):118–20.
  75. Rottingen JA, Cameron DW, Garnett GP. A systematic review of the epidemiologic interactions between classic sexually transmitted diseases and HIV:

- how much really is known? *Sex Transm Dis.* 2001/11/02. 2001;28(10):579–97.
76. Freeman EE, Weiss HA, Glynn JR, Cross PL, Whitworth JA, Hayes RJ. Herpes simplex virus 2 infection increases HIV acquisition in men and women: systematic review and meta-analysis of longitudinal studies. *Aids.* 2006;20(1):73–83.
  77. Cohen MS. HIV and sexually transmitted diseases: lethal synergy. *Top HIV Med.* 2004;12(4):104–7.
  78. Auvert B, Buve A, Ferry B, Carael M, Morison L, Lagarde E, et al. Ecological and individual level analysis of risk factors for HIV infection in four urban populations in sub-Saharan Africa with different levels of HIV infection. *AIDS.* 2001;15:S15–30.
  79. Zhu J, Hladik F, Woodward A, Klock A, Peng T, Johnston C, et al. Persistence of HIV-1 receptor-positive cells after HSV-2 reactivation is a potential mechanism for increased HIV-1 acquisition. *Nat Publ Gr.* 2009;5(5):1–3.
  80. Kaul R, Pettengell C, Sheth PM, Sunderji S, Biringner A, Macdonald K, et al. The genital tract immune milieu: an important determinant of HIV susceptibility and secondary transmission. *J Reprod Immunol.* 2008;77:32–40.
  81. Weiss HA, Buve A, Robinson NJ, Van Dyck E, Kahindo M, Anagonou S, et al. The epidemiology of HSV-2 infection and its association with HIV infection in four urban African populations. *AIDS.* 2001;15(Suppl 4):S97–108.
  82. Kraut-Becher JR, Aral S o. Patterns of age mixing and sexually transmitted infections. *Int J STD AIDS.* 2006;17:378–83.
  83. O’Leary A, Jemmott III JB, Jemmott LS, Teitelman A, Heeren GA, Ngwane Z, et al. Associations between Psychosocial Factors and Incidence of Sexually Transmitted Disease Among South African Adolescents. *Sex Transm Dis.* 2015;42(3):135–9.
  84. Polis CB, Westreich D, Balkus JE, Heffron R, Participants of the HCHIVOAM. Assessing the effect of hormonal contraception on HIV acquisition in observational data: challenges and recommended analytic approaches. *AIDS.* 2013/10/23. 2013;27 Suppl 1:S35-43.
  85. Heffron R, Donnell D, Rees H, Celum C, Mugo N, Were E, et al. Use of hormonal contraceptives and risk of HIV-1 transmission: a prospective cohort study. *Lancet Infect Dis.* 2012;12:19–26.
  86. Butler AR, Smith JA, Polis CB, Gregson S, Stanton D, Hallett TB. Modelling the global competing risks of a potential interaction between injectable hormonal contraception and HIV risk. *AIDS.* 2012/09/28. 2013;27(1):105–13.
  87. Morrison CS, Skoler-Karpoﬀ S, Kwok C, Chen P-L, van de Wijgert J, Gehret-Plagianos M, et al. Hormonal contraception and the risk of HIV acquisition among women in South Africa. *Aids.* 2012;26(4):497–504.
  88. Hawkins K, Price N, Mussa F. Milking the cow: Young women’s construction of identity and risk in age-disparate transactional sexual relationships in Maputo, Mozambique. *Glob Public Heal An Int J Res Policy Pract.* 2012;4(2):169–82.

89. Luke N. Age and economic asymmetries in the sexual relationships of adolescent girls in sub-Saharan Africa. *Stud Fam Plann.* 2003/08/02. 2003;34(2):67–86.
90. Groes-Green C. ‘To put men in a bottle’: Eroticism, kinship, female power, and transactional sex in Maputo, Mozambique. *Am Ethnol.* 2013;40(1):102–17.
91. Moore AM, Biddlecom AE, Zulu EM. Prevalence and meanings of exchange of money or gifts for sex in unmarried adolescent sexual relationships in sub-Saharan Africa. *Afr J Reprod Health.* 2008/05/07. 2007;11(3):44–61.
92. Weiser SD, Leiter K, Bangsberg DR, Butler LM, Percy-de Korte F, Hlanze Z, et al. Food insufficiency is associated with high-risk sexual behavior among women in Botswana and Swaziland. *PLoS Med.* 2007/10/26. 2007;4(10):1589–97; discussion 1598.
93. Nkosana J, Rosenthal D. The dynamics of intergenerational sexual relationships: the experience of schoolgirls in Botswana. *Sex Heal.* 2007/10/13. 2007;4(3):181–7.
94. Zhang L, Iritani BJ, Luseno W, Hartman S, Rusakaniko S, Hallfors DD. Marital Age Disparity Among Orphaned Young Women and Their Husbands: A Mixed Methods Study in Rural Zimbabwe. *J Prim Prev.* 2016;37:487–92.
95. Hawkins K, Price N, Mussa F. Milking the cow: young women’s construction of identity and risk in age-disparate transactional sexual relationships in Maputo, Mozambique. *Glob Public Health.* 2009;4(2):169–82.
96. Kaufman MR, Tsang SW, Mooney A, McCartney-Melstad A, Mushi AK, Kamala B. Protect Your Loved Ones From Fataki? *Qual Health Res.* 2016;26(7):994–1004.
97. Underwood C, Skinner J, Osman N, Schwandt H. Structural determinants of adolescent girls’ vulnerability to HIV: Views from community members in Botswana, Malawi, and Mozambique. *Soc Sci Med.* 2011;73(2):343–50.
98. d’Albis H, Augeraud-Veron E, Djemai E, Ducrot A. The dispersion of age differences between partners and the asymptotic dynamics of the HIV epidemic. *J Biol Dyn.* 2012;6(2):695–717.
99. Bershteyn A, Klein DJ, Eckhoff PA. Age-dependent partnering and the HIV transmission chain: a microsimulation analysis. *J R Soc Interface.* 2013;10(88):20130613.
100. Wellings K, Collumbien M, Slaymaker E, Singh S, Hodges Z, Patel D, et al. Sexual behaviour in context: a global perspective. *Lancet.* 2006/11/14. 2006;368(9548):1706–28.
101. Kenyon CR, Vu L, Menten J, Maughan-Brown B. Male Circumcision and Sexual Risk Behaviors May Contribute to Considerable Ethnic Disparities in HIV Prevalence in Kenya: An Ecological Analysis. *PLoS One.* 2014;9(8):e106230.
102. Robinson WS. Ecological correlations and the behavior of individuals. *Int J Epidemiol.* 2011;40(4):1134.
103. Evans M, Risher K, Zungu N, Shisana O, Moyo S, Celentano DD, et al. Age-disparate sex and HIV risk for young women from 2002 to 2012 in South Africa. *J Int AIDS Soc.* 2016;19(1):1–16.

104. Kaiser R, Bunnell R, Hightower A, Kim AA, Cherutich P, Mwangi M, et al. Factors associated with HIV infection in married or cohabitating couples in Kenya: results from a nationally representative study. *PLoS One*. 2011;6(3):e17842.
105. Schroder KEE, Carey MP, Vanable P a. Methodological challenges in research on sexual risk behavior: II. Accuracy of self-reports. *Ann Behav Med*. 2003;26:104–23.
106. Crawford M, Popp D. Sexual double standards: a review and methodological critique of two decades of research. *J Sex Res*. 2003/06/14. 2003;40(1):13–26.
107. 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. *J Acquir Immune Defic Syndr*. 2003/03/18. 2003;32(4):446–51.
108. Harling G, Newell ML, Tanser F, Kawachi I, Subramanian S V, Barnighausen T. Do Age-Disparate Relationships Drive HIV Incidence in Young Women? Evidence from a Population Cohort in Rural KwaZulu-Natal, South Africa. *J Acquir Immune Defic Syndr*. 2014;66(4):443–51.
109. Harling G, Newell ML, Tanser F, Barnighausen T. Partner Age-Disparity and HIV Incidence Risk for Older Women in Rural South Africa. *AIDS Behav*. 2015;
110. Balkus JE, Nair G, Montgomery ET, Mishra A, Palanee-phillips T, Ramjee G, et al. Age-disparate partnerships and risk of HIV-1 acquisition among South African women participating in the VOICE trial. *J Acquir Immune Defic Syndr*. 2015;70(2):212–7.
111. Schaefer R, Gregson S, Eaton JW, Mugurungi O, Rhead R, Takaruzza A, et al. Age-disparate relationships and HIV incidence in adolescent girls and young women: evidence from a general-population cohort in Zimbabwe. *AIDS*. 2017;
112. de Oliveira T, Kharsany ABM, Gräf T, Cawood C, Khanyile D, Grobler A, et al. Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: a community-wide phylogenetic study. *Lancet HIV*. 2016;3018(16):1–10.
113. Zaidi J, Grapsa E, Tanser F, Newell ML, Barnighausen T. Dramatic increase in HIV prevalence after scale-up of antiretroviral treatment. *AIDS*. 2013;27(14):2301–5.
114. Evans M, Maughan-Brown B, Zungu N, George G. HIV Prevalence and ART Use Among Men in Partnerships with 15–29 Year Old Women in South Africa: HIV Risk Implications for Young Women in Age-Disparate Partnerships. *AIDS Behav*. 2017;
115. Casterline JB, Williams L, Mcdonald P. Population Studies : A Journal of The Age Difference Between Spouses : Variations among Developing Countries The Age Difference Between Spouses " Variations among Developing Countries \*. (March 2013):37–41.
116. Kuate-Defo B. Young People's Relationships with Sugar Daddies and Young People ' s Relationships Sugar Mummies : What do We Know and What do We Need to Know ? *Afr J Reprod Health*. 2008;8(2):13–37.

117. Beauclair R, Meng F, Deprez N, Temmerman M, Welte A, Hens N, et al. Evaluating Audio Computer Assisted Self-Interviews in Urban South African Communities: Evidence for Good Suitability and Reduced Social Desirability Bias of a Cross-sectional Survey on Sexual Behaviour. *BMC Med Res Methodol*. 2013;13(11):11.
118. Phillips AE, Gomez GB, Boily MC, Garnett GP. A systematic review and meta-analysis of quantitative interviewing tools to investigate self-reported HIV and STI associated behaviours in low- and middle-income countries. *Int J Epidemiol*. 2010/07/16. 2010;39(6):1541–55.
119. Dolezal C, Marhefka SL, Santamaria EK, Leu CS, Brackis-Cott E, Mellins CA. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Arch Sex Behav*. 2012;41(2):401–10.
120. Kissinger P, Rice J, Farley T, Trim S, Jewitt K, Margavio V, et al. Application of computer-assisted interviews to sexual behavior research. *Am J Epidemiol*. 1999/05/26. 1999;149(10):950–4.
121. Hewett PC, Mensch BS, Erulkar AS. Consistency in the reporting of sexual behaviour by adolescent girls in Kenya: a comparison of interviewing methods. *Sex Transm Infect*. 2004;80(suppl 2):ii43-ii48.
122. Lam D, Ardington C, Branson N, Case A, Leibbrandt M, Menendez A, et al. The Cape Area Panel Study: A Very Short Introduction to the Integrated Waves 1-2-3-4 Data. Cape Town; 2008.
123. Town C of C. Socio-economic profile: City of Cape Town 2016. Cape Town; 2017.
124. Myer L, Ehrlich RI, Susser ES. Social epidemiology in South Africa. *Epidemiol Rev*. 2004/07/06. 2004;26:112–23.
125. Department SDI and G. City of Cape Town – 2011 Census – Cape Town. Cape Town; 2012.
126. P G, W M, E M, N Z, I N, J D, et al. Western Cape Mortality Profile 2011. Cape Town; 2014.
127. Helleringer S, Mkandawire J, Kalilani-Phiri L, Kohler HP. Cohort Profile: The Likoma Network Study (LNS). *Int J Epidemiol*. 2013/04/02. 2013;
128. Helleringer S, Kohler HP, Chimbiri A, Chatonda P, Mkandawire J. The Likoma Network Study: Context, data collection, and initial results. *Demogr Res*. 2009;21:427–68.
129. Helleringer S, Chimbiri A, Mkandawire J, Kohler H. The Likoma Network Study : Context , Data Collection and Initial Results. 2007;
130. Helleringer S, Kohler HP. Social networks, perceptions of risk, and changing attitudes towards HIV/AIDS: new evidence from a longitudinal study using fixed-effects analysis. *Popul Stud*. 2005;59(3):265–82.
131. Baird SJ, Garfein RS, McIntosh CT, Ozler B. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. *Lancet*. 2012/02/22. 2012;379(9823):1320–9.

132. World Bank T. Impact Evaluation Surveys [Internet]. Central Microdata Catalog. 2017 [cited 2015 Mar 2]. Available from: [http://microdata.worldbank.org/index.php/catalog/impact\\_evaluation](http://microdata.worldbank.org/index.php/catalog/impact_evaluation)
133. Baird S, Chirwa E, Mcintosh C, Özler B. What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi, 3ie Impact Evaluation Report 27. New Delhi; 2015.
134. Baird S, Mcintosh C, Özler B. When the money runs out: Do cash transfers have sustained effects? Unpubl Manusc. 2016;
135. Mpoola L., Chanza C, Nankuyu F, Kamela H, Kaudna L., Chirambo A, et al. Malawi: Zomba Urban Profile. 2011.
136. Baird S, McIntosh C, Özler B. Cash or condition? Evidence from a cash transfer experiment. *Q J Econ.* 2011;126(4):1709–53.
137. Ayles HM, Sismanidis C, Beyers N, Hayes RJ, Godfrey-Faussett P. ZAMSTAR, The Zambia South Africa TB and HIV Reduction study: Design of a 2 x2 factorial community randomized trial. *Trials.* 2008;9.
138. Delva W, Beauclair R, Welte A, Vansteelandt S, Hens N, Aerts M, et al. Age-disparity, sexual connectedness and HIV infection in disadvantaged communities around Cape Town, South Africa: a study protocol. *BMC Public Health.* 2011/08/04. 2011;11(1):616.
139. Delva W, Meng F, Beauclair R, Deprez N, Temmerman M, Welte A, et al. Coital frequency and condom use in monogamous and concurrent sexual relationships in Cape Town, South Africa. *J Int AIDS Soc.* 2013;16:18034.
140. Sefali P. Khayelitsha turns 30. *GroundUp.* 2016;
141. Delft, Cape Town [Internet]. Revolvly. [cited 2017 Jul 4]. Available from: [https://www.revolvly.com/main/index.php?s=Delft, Cape Town&item\\_type=topic](https://www.revolvly.com/main/index.php?s=Delft,CapeTown&item_type=topic)
142. Muzondo IF, Barry M, Dewar D, Whittal J, Africa S, Muzondo IF, et al. Land Conflicts in Informal Settlements: Wallacedene in Cape Town , South Africa. Athens, Greece; 2004. (FIG Working Week 2004).
143. Frith A. Khayelitsha, Census 2011. 2011.
144. Frith A. Delft, Census 2011. 2011.
145. Frith A. Wallacedene, Census 2011 [Internet]. 2011 [cited 2017 Jul 4]. Available from: <https://census2011.adrianfrith.com/place/199018019>
146. Town C of C. City of Cape Town-Census 2001 [Internet]. Vol. 2013. Cape Town: Strategic Development Information ; 2001. Available from: <http://www.capetown.gov.za/en/stats/2001census/Documents/>
147. Harrell Jr. FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Second. Switzerland: Springer Series in Statistics; 2015.
148. Gelman A. Data analysis using regression and multilevel/hierarchical models. New York: Cambridge University Press; 2007.

149. Bates DM. *lme4: Mixed-effects modelling with R*. Springer; 2010.
150. Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, et al. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol*. 2009;24(3).
151. West RM. Generalised Additive Models. In: Tu YK, Greenwood D, editors. *Modern Methods for Epidemiology*. Dordrecht: Springer; 2012. p. 261–78.
152. Royston P. Regression using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. *J R Stat Soc Ser C-Applied Stat*. 1994;43(3):429–67.
153. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med*. 2005;25(1):127–41.
154. Irwin JR, McClelland GH. Negative Consequences of Dichotomizing Continuous Predictor Variables. *J Mark Res*. 2003;40(3):366–71.
155. Altman DG, Royston P. The cost of dichotomising continuous variables. *Br Med J*. 2006;332(7549):1080.
156. de Boor C. *A practical guide to splines*. New York: Springer; 1978.
157. Simon N. Wood. *Generalized Additive Models: An Introduction with R*. Taylor & Francis; 2006.
158. Wood S. *Generalized Additive Models: An Introduction with R*. 2nd ed. Boca Raton: CRC Press; 2017.
159. Hastie TJ, Tibshirani R. *Generalised additive models*. Boca Raton: Chapman and Hall/CRC; 1990.
160. Rubin DB. Inference and missing data. *Biometrika*. 1976;63(3):581–92.
161. van Buuren S. *Flexible Imputation of Missing Data*. Boca Raton: CRC Press; 2012.
162. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4):377–99.
163. Rubin D. *Multiple Imputation for Nonresponse in Surveys*. New York: Wiley; 1987.
164. van Buuren S, Groothuis-Oudshoorn K, Buuren S, Groothuis-Oudshoorn K. MICE: Multivariate Imputation by Chained Equations in R. *J Stat Softw*. 2011;45(3):1–67.
165. Breiman L. Random Forests. *Mach Learn*. 2001;45(1):5–32.
166. Penone C, Davidson AD, Shoemaker KT, DiMarco M, Rondinini C, Brooks TM, et al. Imputation of missing data in life-history trait datasets: which approach performs best? *Methods Ecol Evol*. 2014;5(9):961–70.
167. Shah AD, Bartlett JD, Carpenter J, Nicholas O, Hemingway H. Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study. *Am J Epidemiol*. 2014;179(6):764–74.

168. Stekhoven DJ, Buhlmann P. MissForest - nonparametric missing value imputation for mixed-type data. *Bioinformatics*. 2012;28(1):112–8.
169. Graham JW, Olchowski AE, Gilreath TD. How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. 2007;206–13.
170. Green J, Thorogood N. *Qualitative Methods for Health Research*. London: SAGE Publications Ltd; 2004.
171. Javadi M, Zarea K. Understanding Thematic Analysis and its Pitfall. *J Client Care*. 2016;1(1).
172. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77–101.
173. Saldana J. *The coding manual for qualitative researchers*. CA: Sage Publications; 2015.
174. Boyatzis RE. *Transforming qualitative information: Thematic analysis and code development*. Sage; 1998.
175. Polacek M, Celentano D, O'Campo P, Santelli J. Correlates of condom use stage of change: implications for intervention. *AIDS Educ Prev*. 1999;11(1):38–52.
176. Dunkle K, Stephenson R, Karita E, Chomba E, Kayitenkore K, Vwalika M, et al. New heterosexually transmitted HIV infections in married or cohabiting couples in urban Zambia and Rwanda: an analysis of survey and clinical data. *Lancet*. 2008;371(9631):2183–91.
177. Diekmann O, Heesterbeek JA., Metz JAJ. On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *J Math Biol*. 1990;28(4):365–82.
178. Morris M, Kretzschmar M. Concurrent partnerships and the spread of HIV. *AIDS*. 1997/04/01. 1997;11(5):641–8.
179. Moody J. The importance of relationship timing for diffusion. *Soc Forces*. 2002;81:25–56.
180. Goodreau SM, Cassels S, Kasprzyk D, Montano DE, Greek A, Morris M. Concurrent partnerships, acute infection and HIV epidemic dynamics among young adults in Zimbabwe. *AIDS Behav*. 2010/12/31. 2012;16(2):312–22.
181. Smith J, Nyamukapa C, Gregson S, Lewis J, Magutshwa S, Schumacher C, et al. The Distribution of Sex Acts and Condom Use within Partnerships in a Rural Sub-Saharan African Population. *PLoS One*. 2014;9(2):e88378.
182. Ott MQ, Barnighausen T, Tanser F, Lurie M, Newell ML. Age-gaps in sexual partnerships: seeing beyond 'sugar daddies'. *AIDS*. 2011;25:861–9.
183. Leclerc-Madlala S. HIV/AIDS and the importance of sexual culture and context. *Soc Dyn*. 2002;28:20–41.
184. Cockcroft A, Kunda JL, Kgakole L, Masisi M, Laetsang D, Ho-Foster A, et al. Community views of inter-generational sex: findings from focus groups in

- Botswana, Namibia and Swaziland. *Psychol Health Med.* 2010/09/14. 2010;15(5):507–14.
185. Jewkes R, Dunkle K, Nduna M, Levin J, Jama N, Khuzwayo N, et al. Factors associated with HIV sero-status in young rural South African women: Connections between intimate partner violence and HIV. *Int J Epidemiol.* 2006;35(6):1461–8.
  186. Faergemann C, Lauritsen JM, Brink O, Mortensen PB. Do repeat victims of interpersonal violence have different demographic and socioeconomic characters from non-repeat victims of interpersonal violence and the general population? A population-based case-control study. *Scand J Public Health.* 2010;38(5):524–32.
  187. Volpe EM, Hardie TL, Cerulli C, Sommers MS, Morrison-Beedy D. What's age got to do with it? Partner age difference, power, intimate partner violence, and sexual risk in urban adolescents. *J Interpers Violence.* 2013;28(10):2068–87.
  188. Pandey GK, Dutt D, Banerjee B. Partner and relationship factors in domestic violence: perspectives of women from a slum in Calcutta, India. *J Interpers Violence.* 2009;24(7):1175–91.
  189. Roberts T, Auinger P, Klein J. Predictors of partner abuse in a nationally representative sample of adolescents involved in heterosexual dating relationships. *Violence Vict.* 2006;21(1):81–90.
  190. Hinden MJ, Kishor S, Ansara DL, Inc MI. Intimate Partner Violence among Couples in 10 DHS Countries: Predictors and Health Outcomes. DHS Analytical Studies. Calverton, Maryland, USA; 2008.
  191. Dunbar MS, Kang Dufour MS, Lambdin B, Mudekunya-Mahaka I, Nhamo D, Padian NS. The SHAZ! Project: Results from a Pilot Randomized Trial of a Structural Intervention to Prevent HIV among Adolescent Women in Zimbabwe. *PLoS One.* 2014;9(11):e113621.
  192. Dion Hallfors D, Cho H, Rusakaniko S, Mapfumo J, Iritani B, Zhang L, et al. The Impact of School Subsidies on HIV-Related Outcomes among Adolescent Female Orphans.
  193. Pettifor A, MacPhail C, Selin A, Gómez-Olivé FX, Rosenberg M, Wagner RG, et al. HPTN 068: A Randomized Control Trial of a Conditional Cash Transfer to Reduce HIV Infection in Young Women in South Africa---Study Design and Baseline Results. *AIDS Behav.* 2016;1–20.
  194. Hargreaves JR, Morison L a, Kim JC, Bonell CP, Porter JDH, Watts C, et al. The association between school attendance, HIV infection and sexual behaviour among young people in rural South Africa. *J Epidemiol Community Health.* 2008;62(2):113–9.
  195. Stroeken K, Remes P, De Koker P, Michielsen K, Van Vossolle A, Temmerman M. HIV among out-of-school youth in Eastern and Southern Africa: a review. *AIDS Care.* 2011/07/26. 2011;
  196. Stoner MCD, Edwards JK, Miller WC, Aiello AE, Halpern CT, Julien A, et al. The effect of schooling on age-disparate relationships and number of sexual partners among young women in rural South Africa enrolled in HPTN 068. *JAIDS J Acquir*

Immune Defic Syndr. 2017 Sep;1.

197. Cluver L, Boyes M, Orkin M, Pantelic M, Molwena T, Sherr L. Child-focused state cash transfers and adolescent risk of HIV infection in South Africa: a propensity-score-matched case-control study. *Lancet Glob Heal*. 2013;1(6):e362–70.
198. Rosenberg M, Pettifor A, Twine R, Hughes J, Gomez-Olive FX, Wagner RG, Selin A, MacPhail C KK. Evidence for selection effect and Hawthorne effect in behavioural HIV prevention trial among young women in rural South Africa [Internet]. Available from: <https://www.hptn.org/research/publications/688>
199. Parsons HM. What Happened at Hawthorne? *Science* (80- ). 1974;183(4128):922–32.
200. Haushofer, J., & Shapiro J. The short-term impact of unconditional cash transfers to the poor: Experimental Evidence from Kenya. *Q J Econ*. 2016;131(4):1973–2042.
201. Bianchi M, Bobba M. Liquidity, risk, and occupational choices. *Rev Econ Stud*. 2013;80(2):491–511.
202. Weinhardt LS, Forsyth AD, Carey MP, Jaworski BC, Durant LE. Reliability and Validity of Self-Report Measures of HIV-Related Sexual Behavior: Progress Since 1990 and Recommendations for Research and Practice. *Arch Sex Behav*. 1998;27(2).
203. Coughlin SS. Recall bias in epidemiologic studies. *J Clin Epidemiol*. 1990;43(1):87–91.
204. Westhoff CL. Epidemiologic studies: pitfalls in interpretation. *Dialogues Contracept*. 1995/01/01. 1995;4(5):5–6,8.
205. Gallo MF, Behets FM, Steiner MJ, Hobbs MM, Hoke TH, Van Damme K, et al. Prostate-specific antigen to ascertain reliability of self-reported coital exposure to semen. *Sex Transm Dis*. 2006/07/26. 2006;33(8):476–9.
206. Dolezal C, Marhefka SL, Santamaria EK, Leu CS, Brackis-Cott E, Mellins CA. A comparison of audio computer-assisted self-interviews to face-to-face interviews of sexual behavior among perinatally HIV-exposed youth. *Arch Sex Behav*. 2011/05/24. 2012;41(2):401–10.
207. Kabiru CW, Luke N, Izugbara CO, Zulu E. The Correlates of HIV Testing and Impacts on Sexual Behaviour: evidence from a life history study of young people in Kisumu, Kenya. *BMC Public Health*. 2010;10.
208. Luke N, Clark S, Zulu EM. The relationship history calendar: improving the scope and quality of data on youth sexual behavior. *Demography*. 2011/07/07. 2011;48(3):1151–76.
209. Hart TC, Rennison CM, Gibson C. Revisiting Respondent ‘Fatigue Bias’ in the National Crime Victimization Survey. *J Quant Criminol*. 2005;21(3):345–63.
210. Hobbs, Siegel JS, Swanson DA. Age and Sex Composition. 2nd ed. *The Methods and Materials of Demography*. Amsterdam: Elsevier Academic Press; 2004.
211. Bello Y. Error Detection in Outpatients’ Age Data Using Demographic

- Techniques. *Int J Pure Appl Sci Technol.* 2012;10(1):27–36.
212. Denic S, Khatib F, Saadi H. Quality of age data in patients from developing countries. *J Public Heal.* 2004;26(2):168–71.
  213. Haandrikman K, Rajeswari N V, Hutter I, Ramesh BM. Coping with time - Using a local time-path calendar to reduce heaping in durations. *Time Soc.* 2004;13(2–3):339–62.
  214. Pardeshi GS. Age heaping and accuracy of age data collected during a community survey in the yavatmal district, maharashtra. *Indian J Community Med.* 2010;35(3):391–5.
  215. Pullum TW, Inc. MI. *An Assessment of Age and Date Reporting in the DHS Surveys, 1985-2003.* Calverton, Maryland; 2006.
  216. Helleringer S, Kohler HP, Mkandawire J. Women underestimate the age of their partners during survey interviews: implications for HIV risk associated with age mixing in northern Malawi. *Sex Transm Dis.* 2011;38(11):1030–5.
  217. Harling G, Tanser F, Mutevedzi T, Barnighausen T. Assessing the validity of respondents' reports of their partner's ages in a rural South African population-based cohort. *BMJ Open.* 2015;4:e005638.
  218. Schwartzlander B, Stover J, Hallett T, Atun R, Avila C, Gouws E, et al. Towards an improved investment approach for an effective response to HIV/AIDS. *Lancet.* 2011/06/07. 2011;377(9782):2031–41.
  219. MacQueen KM. Framing the social in biomedical HIV prevention trials: a 20-year retrospective. *J Int AIDS Soc.* 2011;14(2):S3.
  220. Coates TJ, Richter L, Caceres C. Behavioural strategies to reduce HIV transmission: how to make them work better. *Lancet.* 2008;372(9639):669–84.
  221. Lurie MN, Rosenthal S. The Concurrency Hypothesis in Sub-Saharan Africa: Convincing Empirical Evidence is Still Lacking. Response to Mah and Halperin, Epstein, and Morris. *AIDS Behav.* 2010;14:34–7.
  222. Morris M, Vu L, Leslie-Cook A, Akom E, Stephen A, Sherard D. Comparing Estimates of Multiple and Concurrent Partnerships Across Population Based Surveys: Implications for Combination HIV Prevention. *AIDS Behav.* 2014;18(4):783–90.
  223. Akullian A, Bershteyn A, Klein D, Vandormael A, Barnighausen T, Tanser F. Sexual partnership age-pairings and risk of HIV acquisition in rural South Africa. *AIDS.* 2017 Jun;31:1755–64.
  224. Resseguier N, Giorgi R, Paoletti X. Sensitivity Analysis When Data Are Missing Not-at-random. *Epidemiology.* 2011;22(2):279.
  225. Hope R. *Addressing Cross-Generational Sex: A desk review of research and programs.* Population Reference Bureau; 2007.
  226. eNCA. Health department on campaign to warn young girls about sugar daddies. 2016.

227. Zille H. SA Today: A Mind-Shift is needed to win the war against HIV and AIDS [Internet]. 2014 [cited 2017 Aug 2]. Available from: <https://www.da.org.za/2014/08/sa-today-mind-shift-needed-win-war-hiv-aids/>
228. Geffen N. Are sugar daddies bad for your health? GroundUp. 2014;
229. Maughan-Brown B, Venkataramani AS. Accuracy and determinants of perceived HIV risk among young women in South Africa. *BMC Public Health*. 2018;18(1):42.
230. Hill A, Pozniak A. HIV treatment cascades: how can all countries reach the UNAIDS 90-90-90 target? *AIDS*. 2015;29(18):2523–5.
231. UNAIDS. 90-90-90: An ambitious treatment target to help end the AIDS epidemic. 2014.
232. Leclerc-Madlala S. Transactional Modernity, Sex and the Pursuit of Modernity. *Soc Dyn*. 2004;2:1–21.
233. Fox AM, Jackson SS, Hansen NB, Gasa N, Crewe M, Sikkema KJ. In their own voices: A qualitative study of women’s risk for intimate partner violence and HIV in South Africa. *Violence Against Woman*. 2007;13:583–602.
234. Ward CL, van der Merwe A, Dawes A. Youth Violence: Sources and Solutions in South Africa. *Crime and justice: a review of research*,. Cape Town, South Africa: UCT Press; 2012.
235. 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*. 2005/10/20. 2005;2(11):e298.
236. Bailey RC, Moses S, Parker CB, Agot K, Maclean I, Krieger JN, et al. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. *Lancet*. 2007/02/27. 2007;369(9562):643–56.
237. Gray RH, Kigozi G, Serwadda D, Makumbi F, Watya S, Nalugoda F, et al. Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. *Lancet*. 2007/02/27. 2007;369(9562):657–66.
238. Sgaier SK, Reed JB, Thomas A, Njeuhmeli E. Achieving the HIV Prevention Impact of Voluntary Medical Male Circumcision : Lessons and Challenges for Managing Programs. 2014;11(5).