SEXUAL BEHAVIOR AND SEXUAL NETWORKS IN SOUTH AFRICA: IMPLICATIONS FOR HIV TRANSMISSION

by
Kathryn A. Risher, M.H.S.

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

June, 2017

ABSTRACT

HIV is largely spread through sexual transmission in sub-Saharan Africa. Despite major biomedical innovations in HIV prevention, South Africa continues to bear a disproportionate burden of HIV. This dissertation aims to assess sexual behavior among people living with HIV comparing those on antiretroviral therapy (ART) to those not, describe sexual mixing patterns and number of sexual partners, and characterize sexual networks consistent with limited network data and assess the impact of network structure on disease potential. Throughout this dissertation, we present analyses of the Human Sciences Research Council's 2012 South African National HIV Prevalence, Incidence and Behaviour Survey (SABSSM IV), a nationally representative household based cross-sectional survey.

We first use logistic regression to assess the relationship between ART and sexual behavior among those living with HIV. We find that ART is associated with increased odds of condom use among those living with HIV (aOR>2), but not associated with reporting multiple sexual partners. This aim suggests that people living with HIV not yet on ART in South Africa likely contribute the greatest number of transmissions (both due to sexual behavior and ART reducing infectivity), and reinforces the importance of engaging individuals living with HIV in care.

Next, we use mixing matrices and Newman's assortativity coefficients to describe sexual mixing patterns, and fit a number of count distributions to degree distribution (number of sexual partners in the past year) data. Sexual mixing patterns are strongly assortative in South Africa, with assortativity coefficients for age, race, education, HIV-status, number of sexual partners and ARV status indicating strong assortativity in

household partnerships (>0.6). Number of sexual partners was low (mean in past year = 1.34) but men were 5 times more likely to report 2+ partners in the past year. Our findings suggest that the strongly assortative nature of sexual networks in South Africa could have implications for HIV combination prevention intervention efficacy.

Finally, we use a nonparametric Markov chain Monte Carlo approach to simulate complete sexual networks consistent with mixing patterns and degree distribution data. We assess network characteristics on these consistent networks, and assess the impact of assumptions to balance male and female degree on these networks. We then estimate the impact of network structure on disease transmission. Simulated sexual networks consistent with our limited sexual network data varied little, but were highly dependent on assumptions made to balance male and female degree distributions. Networks with FSW populations had the greatest potential for HIV spread. Network characteristics were associated with potential HIV spread. Our results suggest the importance of capturing highly connected individuals in survey data, as these individuals will play a major role in disease transmission.

Sexual networks have the potential to dramatically influence the impact of HIV combination prevention interventions. This dissertation builds upon a body of work to provide an improved understanding of sexual behavior, sexual mixing and sexual networks within South Africa. These results can be utilized in the development of interventions to predict the potential effect an intervention could have in order to efficiently target interventions to have the greatest impact on HIV burden in South Africa.

THESIS READERS

Shruti H. Mehta, Ph.D. (Advisor)
Professor

Department of Epidemiology

Johns Hopkins Bloomberg School of Public Health

Justin Lessler, Ph.D. (Co-Advisor)
Associate Professor

Department of Epidemiology

Johns Hopkins Bloomberg School of Public Health

Maunank Shah, M.D., Ph.D. (Committee Chair)
Associate Professor

Department of Infectious Diseases

Johns Hopkins School of Medicine

Stéphane Helleringer, Ph.D. (Thesis reader)
Associate Professor

Department of Population, Family and Reproductive Health

Johns Hopkins Bloomberg School of Public Health

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my partner, Peter, for putting up with me through all of the ups and downs of a doctoral program and continually encouraging me throughout. Peter has been my sounding board as I develop ideas, and has sat through more hours of discussing sexual networks than I imagine he ever wanted to. He has been my emotional and practical support as I have completed my masters and doctorate and I am eternally grateful to him for his kindness, generosity and love.

I also would like to thank my thesis committee, Justin Lessler, David Celentano and Shruti Mehta. Justin has been unfailingly generous with his time and willingness to engage in my project, has encouraged me to challenge myself and provided a neverending supply of methodological ideas and suggestions. David has been a wonderful mentor and source of wisdom throughout my project and engagement with the Human Sciences Research Council (HSRC). Shruti has been a supportive sounding board and reality check guiding me through this process, who has made herself available to me at a moment's notice.

This dissertation wouldn't be possible without the support given to me by my collaborators at the HSRC, who were willing to engage me in their research process and incorporate me in their team. Particularly, Olive Shisana, Leickness Simbayi, Thomas Rehle and Sean Jooste guided and facilitated my work.

Over the past years, I've had the opportunity to work with a number of researchers at Hopkins, all of whom have helped shape me scientifically. I sincerely appreciate the guidance of David Dowdy, Maunank Shah, Stef Baral, Chris Beyrer, Cecilia Tomori, and Wendy Davis over the past through years in the numerous projects

we have worked on. I am grateful to both of my additional readers, Maunank Shah and Stéphane Helleringer, for your willingness to engage with my dissertation and serve on my defense exam committee.

I couldn't have finished my dissertation without the emotional support given to me by my friends and the members of the IDD team. From hours spent doing practice oral exams to discussing politics, you've helped to keep me grounded in the program. Particularly, Becca Pierce, Keri Calkins, Allison McFall, Kelly Searle, Shaun Truelove Talia Quandelacy, Megan Wallace, Emily Knapp – you have brought me through.

Finally, I would like to thank my parents, Bruce and Pam, for supporting me throughout my (many) years spent in pursuit of education and helping to instill in me a passion for learning.

The research project whose data were used in this dissertation was mainly supported by the President's Emergency Plan for AIDS Relief (PEPFAR) through the CDC under the terms of 5U2GGH000570. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of CDC. Additional support was also received from the Bill & Melinda Gates Foundation, the South African National AIDS Council (SANAC) and the United Nations Children's Fund (UNICEF). My doctoral work was supported by the National Institutes of Health (T32AI102623, F31AI120780) and the Johns Hopkins Epidemiology Department.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

1.1. Overview of the HIV epidemic in South Africa

South Africa has been greatly impacted by the HIV pandemic, with an estimated HIV prevalence in individuals aged 15-49 of 19.2% in 2015[1]. An estimated 7.0 million people are living with HIV infection in South Africa[1], making South Africa the country with the largest number of people living with HIV worldwide and nearly one-sixth of the burden worldwide[4, 5]. Between 2010 and 2013, UNAIDS estimates that the number of new HIV infections dropped by approximately 22%, though there were still 345,000 new infections in 2013[6]. While South Africa has made great strides to decrease the burden of HIV, including having 3.4 million people on antiretroviral treatment (ART)[5] and increasing coverage of HIV counseling and testing (HCT) to over 9.5 million South Africans in a one-year period during 2014-2015[7], HIV remains a major public health problem[8] and was the single disease responsible for the most deaths in 2012[9].

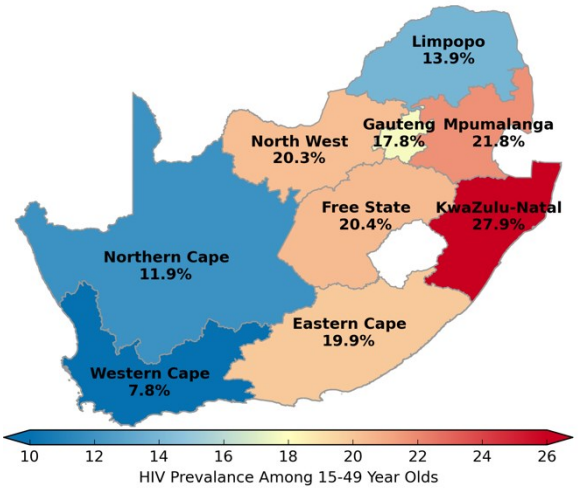


Figure 1.1: Map of South Africa, HIV prevalence among 15-49 year olds by province in 2012 [2]

The HIV epidemic in South Africa has notable heterogeneity along a number of dimensions. HIV in South Africa is geographically extremely disparate (Figure 1.1). HIV prevalence is highest in KwaZulu-Natal (28%), a bit lower in Mpumalanga, Free State, North West, Eastern Cape, and Gauteng provinces (18-22%), and quite substantially

lower in Limpopo, Northern Cape and Western Cape (8-14%). There is also substantial variability by age and sex, shown in Figure 1.2. HIV prevalence is higher among females than males within all age groups. The HIV prevalence among women peaks at ages 30-34 at 36%, while it peaks at ages 35-39 among males at 28.8%. Peak HIV prevalence shifted between 2008 and 2012, from the peak among women being at ages 25-29 and among men being at ages 30-34[10]. This change has also come with an overall increase in HIV prevalence in that period, attributed to the increased availability of ART and consequently improved survival while maintaining high incidence. Sex prevalence differentials in young adults (aged 15-19 and 20-25) has led to a focus on age mixing and the potential role of older male partners in introducing the HIV epidemic to younger women[11]. There are additionally major disparities in the burden of HIV by race in South Africa. Black Africans have the highest HIV prevalence (22.7%), followed by coloured (4.6%), Indian/Asian (1.0%) and white (0.6%)[2].

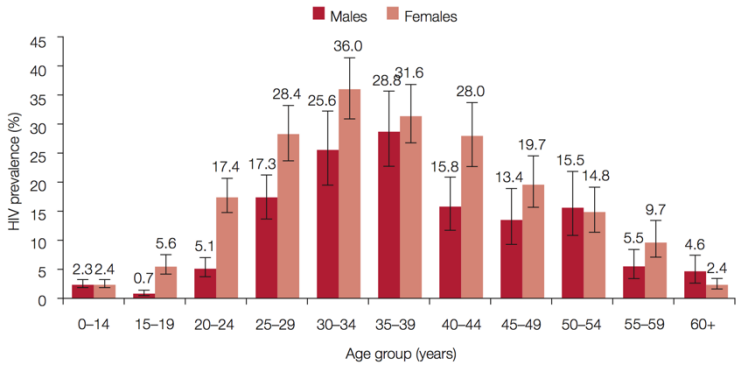


Figure 1.2: HIV prevalence by age and sex in South Africa, 2012[2]

1.2. Sexual behaviors, sexual networks, and HIV

The majority of HIV transmission globally[4] and in Southern Africa[12] occurs via sexual transmission. Sexual transmission has been challenging to elucidate, partly due to the sensitive nature of sexual behavior data, and partly due to the fact that an individual's sexual partners' behaviors can have as much impact on their infection risk as the

individual's own behaviors[13]. As a result, both sexual behaviors[14] and sexual networks[15-18], the networks formed by connecting individuals through their sexual partnerships, have become an increasingly studied area of research.

One measure of individual sexual risk that is measured frequently is number of partners in the past year (referred to as "degree" in network literature). In general populations in Eastern and Southern Africa, the mean number of sexual partners in the past year reported by sexually experienced men ranged from 1.2-1.9 in different sites, and among women ranged from 0.82-1.1[19]. Across studies, men report from 20-100% more sexual partners than women report[20].

Individual sexual behavior is challenging to ascertain accurately, because it is highly personal, and relies almost exclusively on self-report[21]. There has been some work that uses semen biomarker detection from vaginal swabs in heterosexual sex, primarily as a validation of self-report, though this represents a relatively short window of recall (1-2 days) and has not been widely used to ascertain sexual behavior[22]. Previous work has shown sexual behavior reports are strongly influenced by survey methodology [21, 23-26], and has suggested that non-interviewer administered questionnaires (such as audio computer-assisted self-interviewing – ACASI) may provide more valid results.

1.2.1. Sexual behavior & ART status

A recent randomized controlled trial (RCT), HPTN 052, measuring the transmission of HIV from HIV-positive individuals to their HIV-negative (serodiscordant) stable partners showed a 93% decrease in linked transmission events among individuals on immediate ART compared to those who had a delayed start[27,

28]. Results from systematic reviews of observational studies assessing the decrease in HIV transmission among serodiscordant heterosexual partners range from 34-91% reduction in incidence rate, with differences by study population and setting, CD4 count at which ART was initiated, and strength of evidence[29, 30]. A recent observational study found that among serodifferent couples that reported condomless sex acts while the seropositive partner had suppressed viral load there were no phylogenetically linked transmissions[31]. These findings suggest that the sexual behaviors of those living with HIV not on ART are of the greatest importance to HIV transmission. Other investigators suggest that risk compensation (increased sexual risk behavior reported by individuals who feel protected by protective interventions, in this case ART[32]) will possibly negate the protective effects of ART on transmission[33, 34]. Modeling studies suggest that sexual behavior change would have to be quite dramatic for it to overcome the protective nature of ART[35].

There was early evidence suggesting that individuals on ART had greater risk behavior than those living with HIV not yet on ART in Western countries[36-39]. However, literature published in South Africa[40-43] and other settings in sub-Saharan Africa[44-53] (including a meta-analysis published in 2012[54]) found either no significant difference in sexual risk behavior or a decrease in sexual risk behavior between individuals on ART compared to those who are not. Few studies showed support for an increase in sexual risk behavior by individuals on ART[55-57].

1.2.2. Sexual network data collection methods

Collection of sexual network data generally ranges from sociometric data, in which all individuals in a network are captured, and egocentric data, where only an

individual's personal networks are captured, with a number of "partial network" study designs such as snowball sampling or respondent driven sampling, falling in the middle of those[13]. Figure 1.3 visually depicts some types of network data where each dot represents an individual, the lines represent a sexual partnership between those two individuals, and colors indicate the node's sex (this network is exclusively heterosexual). In Figure 1.3a, sociometric data allow us to see the full, true underlying network, by conducting a census of an area, and specifically naming and identifying each of a person's partners within that area from the census list. Figure 1.3b shows egocentric data, where the survey only captures individuals in boxes, but we are able to learn from the sampled individuals about the individuals with whom that person has relationships. We are unable to see individuals who are greyed out. Figure 1.3c shows household data, which can be used to understand mixing between two individuals, but doesn't allow us to know about other partners. Other study designs that allow for inference about sexual networks include molecular biology/phylogenetic data and public health contact tracing data[11, 15].

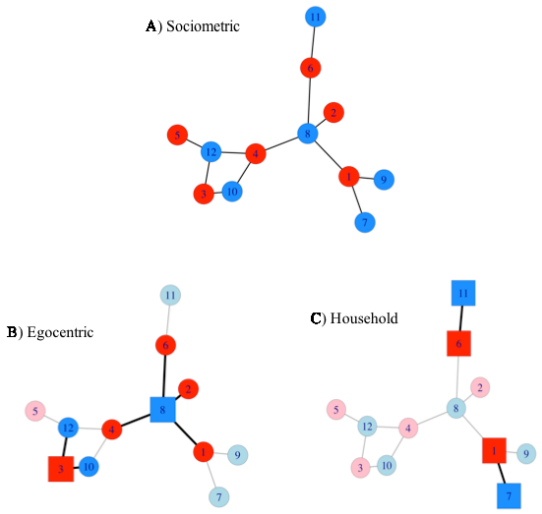


Figure 1.3: Visual depictions of sociometric, egocentric and household network data sources

Each network study design has strengths and weaknesses. Sociometric data provides the most complete picture of a network, but is limited by its high cost and time-intensiveness. Selecting a boundary of the network can be challenging in our highly-connected world[13]. Egocentric data, while relatively easy to collect, does not inform us

about higher order network structures[17]. Additionally, egocentric data has similar validity concerns as individual-level self-reported risk behavior[58, 59].

The most complete source of empirical sexual network data in SSA comes from the Likoma Network Study (LNS) conducted on Likoma Island, Malawi. The authors described sociometric sexual networks in the three years prior to survey and found that the majority (86%) of components (individuals who were linked through sexual partnerships) were of size five or smaller but that these components represented 34% of the respondents, while two-thirds of the population were in 35 (of 256) components of size six or larger[60]. There was a "giant" component of 883 (of 1803) network members, which was characterized by an overall higher average degree (3 vs 1.8) but still had 40% participants who reported 2 or less partners in the three previous years. Notably, in this cross-sectional assessment, the sparser regions of the network had higher HIV prevalence, though there are demographic and risk factors that likely explain this paradoxical observation[60, 61]. The LNS also found that women were more likely to report relationships with visitors, potentially efficient bridges to introduce HIV to communities and their regular partnerships[62].

1.2.3. Sexual network implications for disease transmission

The basic reproductive number (R_0, frequently referred to erroneously as a rate), the average number of secondary infections following the introduction of one infected individual to a fully susceptible population, is designated as the following:

$$R_0 = \beta cd$$

where β is equal to the per contact probability of infection, c is the number of contacts per unit time, and d is the duration of infectiousness[63]. R_0 has implications for the

capability of a disease to spread in a population, if $R_0 > 1$, there is potential for the epidemic to spread throughout the population. If $R_0 < 1$, the disease will ultimately fade out of the population. The epidemic threshold, $1 - \frac{1}{R_0}$, represents the proportion of the population that must be immune to a disease, under assumptions of random mixing, for the disease to be eliminated from the population.

Anderson and May[63] show that average number of contacts can be rewritten as:

$$c = m + \frac{\sigma^2}{m}$$

where m is the mean, and σ^2 is the variance of number of contacts. Thus, the c in the calculation of R_0 for a sexually transmitted disease is not a simple arithmetic mean of number of sexual partners. They emphasize that "superspreaders" who dramatically increase the variance of number of partners have a larger than expected impact on transmission of infection and maintenance of a disease within a population.

Following upon this work, a highly influential paper found that sexual degree distributions in Sweden can be characterized by a scale-free power law[64]. Power laws follow a distribution function:

$$P(k) \approx k^{-\alpha}$$

where k is equal to degree and α is the scaling parameter. When the scaling parameter is $2 < \alpha \leq 3$, the network formed by this degree distribution is referred to as a scale-free network and this degree distribution has infinite variance. Based on this infinite variance, authors find that scale free networks have no epidemic threshold[65-68], though finite population sizes may impact this finding[67]. A number of models simulating scale-free networks found that removing the most connected nodes restores the epidemic threshold[65]. Potential interventions could include targeted provision of combination

HIV prevention to those with the greatest number of partners, such as condoms, ART if living with HIV, and PrEP if HIV-negative. Data from Zimbabwe[69] and among males from Burkina Faso[70] support the scale-free hypothesis. Some teams have criticized the rush of scale-free research[71-74], due to limitations to the methods used to characterize the power-law scaling parameter or concerns about capturing the tails of the distributions, suggesting that these networks may not all be scale-free. However, the strategies to intervene in distributions with highly skewed degree even if they are not scale-free likely remain the same[75].

1.2.4. Assortative Mixing

Assortative mixing, the concept of "like-with-like" partnering[13], is observed on numerous networks[76]. Newman's assortativity coefficient has been used in non-African settings to describe sexual mixing of key populations[77-80]. It was recently used to assess assortativity by educational attainment between partners using two rounds of DHS data from 7 countries in sub Saharan Africa (Cameroon, Ethiopia, Kenya, Lesotho, Malawi, Rwanda, and Zimbabwe), where assortativity coefficients for education ranged from 0.09 in urban Zimbabwe to 0.44 in non-urban Cameroon[81]. One study which assessed sexual mixing by ethnicity in the 1999 Carletonville Youth Survey[82] found that ethnic mixing of sexual partners was homophilous among 15-24 year old participants, though the extent to which this was observed varied by ethnicity under consideration. A recent modeling study found that ART assortativity could impact the population level effect of ART on HIV incidence, but that this effect is modified by the prevalence of HIV, the adherence to ART and the prevalence of ART[83].

Theoretical work has shown that assortativity by degree, or assortativity by "activity class" can influence the shape of an HIV epidemic. Extremely high assortativity by degree is associated with a rapid initial spread of HIV but ultimately a lower overall epidemic, as the epidemic exhausts itself within this "high risk" group of individuals[84, 85]. High assortativity by degree could result in epidemics with multiple peaks across years[85]. Extremely low assortativity by degree is associated with a longer and ultimately more impactful epidemic[84, 85]. Some models of the HIV epidemic in South Africa include sexual activity level mixing, i.e. between low- and high-level sexual risk groups, though these parameters were exclusively calibrated, as opposed to empirically derived[86-89].

1.2.5. Core Group Theory

A classic sexually transmitted infection (STI) theory, core group theory suggests that a core group of individuals who experience the majority of sexually transmitted infection (STI) disease burden and mix predominantly with themselves, keep STIs prevalent by regular sexual activity and partner turnover[90]. These individuals are also hypothesized to be responsible for disease transmission outside of the core group to peripheral groups through occasional non-core-group partnerships. A mathematical modeling study found that the impact of core group mixing is dramatically different by the transmission probability of the infective organism. With lower infectivity per sexual contact assumed for HIV, partner mixing did not have implications for the epidemic size, but with higher transmission probability assumed, increasing partner mixing led to increased epidemic size[91]. Recent work suggests that both core groups with high partner turnover, and the duration that a higher risk individual remains in the "core

setting" (i.e. duration of sex work or duration as a client of sex work) can impact HIV transmission [92].

1.2.6. Concurrency

Concurrent sexual partnerships, or overlapping sexual partnerships with two or more different partners in which sexual acts with one partner occur between sexual acts with another partner[93], have been a topic of extensive debate among HIV researchers for the past two decades[3, 94-101]. Concurrency is theorized to be an important driver of HIV transmission throughout sexual networks because it results in more partnerships which overlap during the acute phase of HIV infection, which is the period immediately after infection that is characterized by high viral load and a higher transmission probability[101]. This is in contrast to serial monogamy, in which earlier partners are not at risk of HIV infection from subsequent partners, and in which the majority of new partnerships will occur after the infected partner has passed from the acute phase into the asymptomatic phase, which is associated with a much lower per act transmission probability[102-104]. A competing hypothesis is that there is coital dilution (i.e. lower number of sex acts) with concurrent partners[105, 106]. There is mixed evidence for the coital dilution hypothesis[104, 107]. The measurement of concurrency has been a

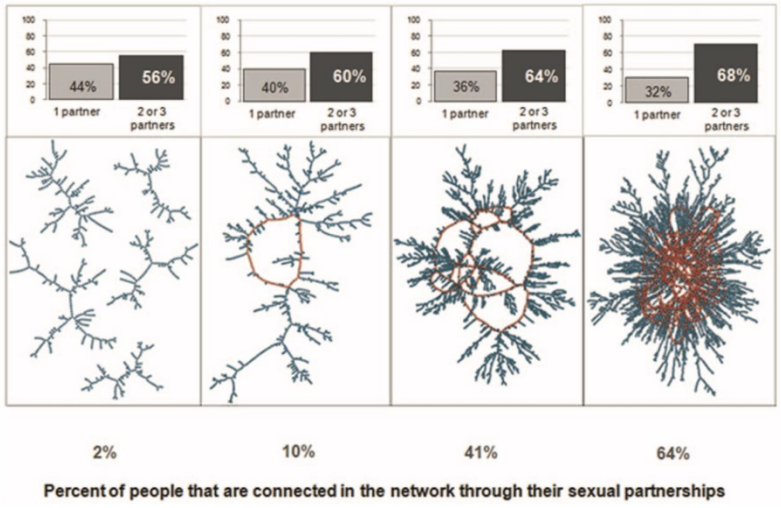


Figure 1.4: Theoretical results on concurrency[3]

challenge for researchers, given the sensitive nature of the sexual behavior data being collected, lack of uniformity in definition across researchers, and issues of temporality [108, 109]. Theoretical results (Figure 1.4) have found that relatively small increases in concurrency can lead to dramatic changes in the connectivity of sexual networks[3].

1.3. Mathematical Models of HIV transmission

Classical mathematical models of disease transmission assume random mixing between individuals constructed using a series of differential equations to describe spread from "infected" to "susceptible" populations[63, 110]. When these models, referred to as compartmental models, are applied to HIV transmission without additional "compartments", they assume that sexual partnerships are selected at random and that partnerships are not ongoing (i.e. long-term partnerships do not exist). Given that these assumptions are controverted by evidence of assortativity in sexual partnership mixing and lengthy duration of live-in and marital partnerships, compartmental models of HIV transmission frequently add population structure (such as high vs low-risk sexual behavior[84], or age structure[111]) to incorporate key sexual network characteristics. Work has shown that differentials in sexual activity by age (high activity in the young, and lower in the older) enhances HIV transmission[84].

In contrast, network models simulate individuals who are connected to one another through sexual partnerships across which HIV transmits[112]. Network models are dramatically more computationally intensive than compartmental models, and involve two simulation steps, first the simulation of the sexual network, followed by the simulation of disease transmission upon this network. Pair formation models have been developed as a middle ground between compartmental and network models[113, 114].

Exponential random graph models (ERGMs) have increasingly been used in mathematical modeling of HIV and STI transmission[115, 116], as they provide a statistical model of sexual networks which can be used as the groundwork for a transmission model. However, most mathematical models that utilize ERGMs (including the limited number of ERGM-based models of the sub-Saharan epidemic[116-118]) do not present the ERGM coefficients from the cross-sectional model that is fit. The presentation of ERGM network statistics would allow other researchers to parameterize and fit sexual networks in the area of interest using widely utilized statistical packages[119-121].

Mathematical models of the HIV epidemic in SSA frequently do not sufficiently incorporate sexual mixing. In a mathematical modeling study compiling ten recent models of HIV transmission in South Africa[122], the models' sexual mixing parameter sources were limited and not clearly cited. Several models incorporated age mixing[123-127], with the parameters coming from the Africa Centre team in rural KwaZulu-Natal in 2011[128], the 1998 South African Demographic and Health Survey (DHS)[129], personal communications on mixing patterns in Guguletu (a township near Cape Town) and Carletonville (near Johannesburg)[123], and unclear source[130]. None of the models incorporate sexual mixing on characteristics other than age and sexual risk level.

1.4. References

1. UNAIDS. South Africa HIV and AIDS estimates. 2015. Available from: http://www.unaids.org/en/regionscountries/countries/southafrica. Accessed 1 Feb 2017.

2. Shisana O, Rehle T, Simbayi L, et al. South African National HIV Prevalence, Incidence and Behaviour Survey, 2012. Cape Town: HSRC Press; 2014.

3. Carnegie NB, Morris M. Size matters: concurrency and the epidemic potential of HIV in small networks. PLoS One. 2012;7(8):e43048.

4. Abdool Karim SS, Abdool Karim Q, Gouws E, Baxter C. Global epidemiology of HIV-AIDS. Infect Dis Clin North Am. 2007;21(1):1-17.

5. UNAIDS. Global AIDS Update. 2016. Available from: http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf. Accessed 1 Feb 2017.

6. UNAIDS. The Gap Report. 2014. Available from: http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2014/UNAIDS_Gap_report_en.pdf. Accessed 14 Aug 2016.

7. South African National AIDS Council. South African Global AIDS Response Progress Report 2015. 2016. Available from: http://sanac.org.za/wp-content/uploads/2016/06/GARPR_report-high-res-for-print-June-15-2016.pdf. Accessed 1 Feb 2017.

8. South African National AIDS Council. Global AIDS Response Progress Report. Republic of South Africa 2012. Available from: http://www.unaids.org/en/dataanalysis/knowyourresponse/countryprogressreports/2012countries/ce_ZA_Narrative_Report.pdf. Accessed

9. Pillay-van Wyk V, Msemburi W, Laubscher R, et al. Mortality trends and differentials in South Africa from 1997 to 2012: second National Burden of Disease Study. Lancet Glob Health. 2016;4(9):e642-53.

10. 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: HSRC Press; 2009.

11. de Oliveira T, Kharsany AB, Graf T, et al. Transmission networks and risk of HIV infection in KwaZulu-Natal, South Africa: a community-wide phylogenetic study. Lancet HIV. 2017;4(1):e41-e50.

12. 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.

13. Morris M, International Union for the Scientific Study of Population. Network epidemiology : a handbook for survey design and data collection. Oxford ; New York: Oxford University Press; 2004. xii, 237 p. p.

14. Wellings K, Collumbien M, Slaymaker E, et al. Sexual behaviour in context: a global perspective. Lancet. 2006;368(9548):1706-28.

15. Delva W, Leventhal GE, Helleringer S. Connecting the dots: network data and models in HIV epidemiology. AIDS. 2016;30(13):2009-20.

16. Liljeros F. Human Sexual Networks. In: Meyers RA, editor. Computational Complexity: Theory, Techniques, and Applications. New York: Springer; 2012.

17. Doherty IA, Padian NS, Marlow C, Aral SO. Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. J Infect Dis. 2005;191 Suppl 1:S42-54.

18. Rothenberg R. HIV transmission networks. Curr Opin HIV AIDS. 2009;4(4):260-5.

19. Todd J, Cremin I, McGrath N, et al. Reported number of sexual partners: comparison of data from four African longitudinal studies. Sex Transm Infect. 2009;85 Suppl 1:i72-80.

20. 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.

21. Cleland J, Boerma JT, Carael M, Weir SS. Monitoring sexual behaviour in general populations: a synthesis of lessons of the past decade. Sex Transm Infect. 2004;80 Suppl 2:ii1-7.

22. Gallo MF, Steiner MJ, Hobbs MM, Warner L, Jamieson DJ, Macaluso M. Biological markers of sexual activity: tools for improving measurement in HIV/sexually transmitted infection prevention research. Sex Transm Dis. 2013;40(6):447-52.

23. 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:ii43-8.

24. Helleringer S, Kohler HP, Kalilani-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.

25. Mensch BS, Hewett PC, Erulkar AS. The reporting of sensitive behavior by adolescents: a methodological experiment in Kenya. Demography. 2003;40(2):247-68.

26. 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.

27. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med. 2011;365(6):493-505.

28. Cohen MS, Chen YQ, McCauley M, et al. Antiretroviral Therapy for the Prevention of HIV-1 Transmission. N Engl J Med. 2016;375(9):830-9.

29. Anglemyer A, Rutherford GW, Horvath T, Baggaley RC, Egger M, Siegfried N. Antiretroviral therapy for prevention of HIV transmission in HIV-discordant couples. Cochrane Database Syst Rev. 2013;4:CD009153.

30. Baggaley RF, White RG, Hollingsworth TD, Boily MC. Heterosexual HIV-1 infectiousness and antiretroviral use: systematic review of prospective studies of discordant couples. Epidemiology. 2013;24(1):110-21.

31. Rodger AJ, Cambiano V, Bruun T, et al. Sexual Activity Without Condoms and Risk of HIV Transmission in Serodifferent Couples When the HIV-Positive Partner Is Using Suppressive Antiretroviral Therapy. JAMA. 2016;316(2):171-81.

32. Cassell MM, Halperin DT, Shelton JD, Stanton D. Risk compensation: the Achilles' heel of innovations in HIV prevention? BMJ. 2006;332(7541):605-7.

33. Wilson DP. HIV treatment as prevention: natural experiments highlight limits of antiretroviral treatment as HIV prevention. PLoS Med. 2012;9(7):e1001231.

34. Venkatesh KK, Flanigan TP, Mayer KH. Is expanded HIV treatment preventing new infections? Impact of antiretroviral therapy on sexual risk behaviors in the developing world. AIDS. 2011;25(16):1939-49.

35. Jean K, Boily MC, Danel C, et al. What Level of Risk Compensation Would Offset the Preventive Effect of Early Antiretroviral Therapy? Simulations From the TEMPRANO Trial. Am J Epidemiol. 2016.

36. Stolte IG, Dukers NH, Geskus RB, Coutinho RA, de Wit JB. Homosexual men change to risky sex when perceiving less threat of HIV/AIDS since availability of highly active antiretroviral therapy: a longitudinal study. AIDS. 2004;18(2):303-9.

37. Dukers NH, Goudsmit J, de Wit JB, Prins M, Weverling GJ, Coutinho RA. Sexual risk behaviour relates to the virological and immunological improvements during highly active antiretroviral therapy in HIV-1 infection. AIDS. 2001;15(3):369-78.

38. Ostrow DE, Fox KJ, Chmiel JS, et al. Attitudes towards highly active antiretroviral therapy are associated with sexual risk taking among HIV-infected and uninfected homosexual men. AIDS. 2002;16(5):775-80.

39. Tun W, Gange SJ, Vlahov D, Strathdee SA, Celentano DD. Increase in sexual risk behavior associated with immunologic response to highly active antiretroviral therapy among HIV-infected injection drug users. Clin Infect Dis. 2004;38(8):1167-74.

40. Eisele TP, Mathews C, Chopra M, et al. High levels of risk behavior among people living with HIV Initiating and waiting to start antiretroviral therapy in Cape Town South Africa. AIDS Behav. 2008;12(4):570-7.

41. Peltzer K, Ramlagan S. Safer sexual behaviours after 1 year of antiretroviral treatment in KwaZulu-Natal, South Africa: a prospective cohort study. Sex Health. 2010;7(2):135-41.

42. Venkatesh KK, de Bruyn G, Lurie MN, et al. Decreased sexual risk behavior in the era of HAART among HIV-infected urban and rural South Africans attending primary care clinics. AIDS. 2010;24(17):2687-96.

43. Eisele TP, Mathews C, Chopra M, et al. Changes in risk behavior among HIV-positive patients during their first year of antiretroviral therapy in Cape Town South Africa. AIDS Behav. 2009;13(6):1097-105.

44. Apondi R, Bunnell R, Ekwaru JP, et al. Sexual behavior and HIV transmission risk of Ugandan adults taking antiretroviral therapy: 3 year follow-up. AIDS. 2011;25(10):1317-27.

45. Dia A, Marcellin F, Bonono RC, et al. Prevalence of unsafe sex with one's steady partner either HIV-negative or of unknown HIV status and associated determinants in Cameroon (EVAL ANRS12-116 survey). Sex Transm Infect. 2010;86(2):148-54.

46. Bateganya M, Colfax G, Shafer LA, et al. Antiretroviral therapy and sexual behavior: a comparative study between antiretroviral- naive and -experienced patients at an urban HIV/AIDS care and research center in Kampala, Uganda. AIDS Patient Care STDS. 2005;19(11):760-8.

47. Moatti JP, Prudhomme J, Traore DC, et al. Access to antiretroviral treatment and sexual behaviours of HIV-infected patients aware of their serostatus in Cote d'Ivoire. AIDS. 2003;17 Suppl 3:S69-77.

48. Luchters S, Sarna A, Geibel S, et al. Safer sexual behaviors after 12 months of antiretroviral treatment in Mombasa, Kenya: a prospective cohort. AIDS Patient Care STDS. 2008;22(7):587-94.

49. Wandera B, Kamya MR, Castelnuovo B, et al. Sexual behaviors over a 3-year period among individuals with advanced HIV/AIDS receiving antiretroviral therapy in an urban HIV clinic in Kampala, Uganda. J Acquir Immune Defic Syndr. 2011;57(1):62-8.

50. Thirumurthy H, Goldstein M, Zivin JG, Habyarimana J, Pop-Eleches C. Behavioral Responses of Patients in AIDS Treatment Programs: Sexual Behavior in Kenya. Forum Health Econ Policy. 2012;15(2).

51. Yalew E, Zegeye DT, Meseret S. Patterns of condom use and associated factors among adult HIV positive clients in North Western Ethiopia: a comparative cross sectional study. BMC Public Health. 2012;12:308.

52. Bunnell R, Ekwaru JP, Solberg P, et al. Changes in sexual behavior and risk of HIV transmission after antiretroviral therapy and prevention interventions in rural Uganda. AIDS. 2006;20(1):85-92.

53. Jean K, Gabillard D, Moh R, et al. Effect of early antiretroviral therapy on sexual behaviors and HIV-1 transmission risk among adults with diverse heterosexual partnership statuses in Cote d'Ivoire. J Infect Dis. 2014;209(3):431-40.

54. Berhan A, Berhan Y. Is the Sexual Behaviour of HIV Patients on Antiretroviral therapy safe or risky in Sub-Saharan Africa? Meta-Analysis and Meta-Regression. AIDS Res Ther. 2012;9(1):14.

55. Olley BO. Higher-risk sexual behaviour among HIV patients receiving antiretroviral treatment in Ibadan, Nigeria. Afr J AIDS Res. 2008;7(1):71-8.

56. Shafer LA, Nsubuga RN, White R, et al. Antiretroviral therapy and sexual behavior in Uganda: a cohort study. AIDS. 2011;25(5):671-8.

57. Pearson CR, Cassels S, Kurth AE, Montoya P, Micek MA, Gloyd SS. Change in sexual activity 12 months after ART initiation among HIV-positive Mozambicans. AIDS Behav. 2011;15(4):778-87.

58. Harling G, Tanser F, Mutevedzi T, Barnighausen T. Assessing the validity of respondents' reports of their partners' ages in a rural South African population-based cohort. BMJ Open. 2015;5(3):e005638.

59. 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.

60. Helleringer S, Kohler HP. Sexual network structure and the spread of HIV in Africa: evidence from Likoma Island, Malawi. AIDS. 2007;21(17):2323-32.

61. Helleringer S, Kohler H-P. Cross-sectional research deisgn and relatively low HIV incidence, rather than blood exposures, explain the peripheral location of HIV cases within the sexual networks observed on Likoma. AIDS. 2008;22(11):1378-9.

62. Helleringer S, Kohler HP, Chimbiri A. Characteristics of external/bridge relationships by partner type and location where sexual relationship took place. AIDS. 2007;21(18):2560-1.

63. Anderson RM, May RM. Infectious diseases of humans : dynamics and control. Oxford ; New York: Oxford University Press; 1991. viii, 757 p. p.

64. Liljeros F, Edling CR, Amaral LA, Stanley EH, Åberg Y. The web of human sexual contacts. Nature. 2001;411(6840):907-8.

65. Dezso Z, Barabasi AL. Halting viruses in scale-free networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2002;65(5 Pt 2):055103.

66. Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. Phys Rev Lett. 2001;86(14):3200-3.

67. May RM, Lloyd AL. Infection dynamics on scale-free networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2001;64(6 Pt 2):066112.

68. Boguna M, Pastor-Satorras R, Vespignani A. Absence of epidemic threshold in scale-free networks with degree correlations. Phys Rev Lett. 2003;90(2):028701.

69. Schneeberger A, Mercer CH, Gregson SA, et al. Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. Sex Transm Dis. 2004;31(6):380-7.

70. Latora V, Nyamba A, Simpore J, et al. Network of sexual contacts and sexually transmitted HIV infection in Burkina Faso. J Med Virol. 2006;78(6):724-9.

71. Hamilton DT, Handcock MS, Morris M. Degree distributions in sexual networks: a framework for evaluating evidence. Sexually transmitted diseases. 2008;35(1):30-40.

72. Handcock MS, Jones JH. Likelihood-based inference for stochastic models of sexual network formation. Theor Popul Biol. 2004;65(4):413-22.

73. Jones JH, Handcock. An assessment of preferential attachment as a mechanism for human sexual network formation. Proceedings of the Royal Society B: Biological Sciences. 2003;270(1520):11231128.

74. Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. arXiv. 2009;0706.1062.

75. Garnett GP. Inferring mechanisms for sexual partnership formation from the distribution of sexual partner numbers. Sex Transm Dis. 2008;35(1):41-2.

76. Newman MEJ. Networks : an introduction. Oxford ; New York: Oxford University Press; 2010. xi, 772 p. p.

77. Bohl DD, McFarland W, Raymond HF. Improved measures of racial mixing among men who have sex with men using Newman's assortativity coefficient. Sexually transmitted infections. 2011;87(7):616-20.

78. Fujimoto K, Williams ML. Racial/Ethnic Differences in Sexual Network Mixing: A Log-Linear Analysis of HIV Status by Partnership and Sexual Behavior Among Most at-Risk MSM. AIDS and behavior. 2015;19(6):996-1004.

79. Doherty IA, Schoenbach VJ, Adimora AA. Sexual Mixing Patterns and Heterosexual HIV Transmission Among African Americans in the Southeastern United States. JAIDS Journal of Acquired Immune Deficiency Syndromes. 2009;52(1):114.

80. Schneider JA, Cornwell B, Ostrow D, et al. Network mixing and network influences most linked to HIV infection and risk behavior in the HIV epidemic among black men who have sex with men. American journal of public health. 2013;103(1):36.

81. Harling G, Barnighausen T. The role of partners' educational attainment in the association between HIV and education amongst women in seven sub-Saharan African countries. J Int AIDS Soc. 2016;19(1):20038.

82. Kenyon C, Colebunders R. Birds of a feather: homophily and sexual network structure in sub-Saharan Africa. International journal of STD & AIDS. 2013;24(3):211-5.

83. Delva W, Helleringer S. Beyond Risk Compensation: Clusters of Antiretroviral Treatment (ART) Users in Sexual Networks Can Modify the Impact of ART on HIV Incidence. PLoS One. 2016;11(9):e0163159.

84. Anderson RM, Gupta S, Ng W. The significance of sexual partner contact networks for the transmission dynamics of HIV. J Acquir Immune Defic Syndr. 1990;3(4):417-29.

85. Gupta S, Anderson RM, May RM. Networks of sexual contacts: implications for the pattern of spread of HIV. AIDS. 1989;3(12):807-17.

86. Dorrington R, Johnson L, Budlender D. ASSA2008 AIDS and demographic models: user guide. Cape Town: ASSA; 2010. Accessed 15 Aug 2015.

87. Eaton JW, Hallett TB. Why the proportion of transmission during early-stage HIV infection does not predict the long-term impact of treatment on HIV incidence. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(45):16202-7.

88. Cori A, Ayles H, Beyers N, et al. HPTN 071 (PopART): a cluster-randomized trial of the population impact of an HIV combination prevention intervention including universal testing and treatment: mathematical model. PloS one. 2014;9(1).

89. Johnson LF, Hallett TB, Rehle TM, Dorrington RE. The effect of changes in condom usage and antiretroviral treatment coverage on human immunodeficiency virus incidence in South Africa: a model-based analysis. Journal of the Royal Society, Interface / the Royal Society. 2012;9(72):1544-54.

90. Yorke JA, Hethcote HW, Nold A. Dynamics and control of the transmission of gonorrhea. Sex Transm Dis. 1978;5(2):51-6.

91. Stigum H, Falck W, Magnus P. The core group revisited: the effect of partner mixing and migration on the spread of gonorrhea, Chlamydia, and HIV. Math Biosci. 1994;120(1):1-23.

92. Watts C, Zimmerman C, Foss AM, Hossain M, Cox A, Vickerman P. Remodelling core group theory: the role of sustaining populations in HIV transmission. Sex Transm Infect. 2010;86 Suppl 3:iii85-92.

93. Unaids Reference Group on Estimates, Modelling, Projections: Working Group on Measuring Concurrent Sexual Partnerships. HIV: consensus indicators are needed for concurrency. Lancet. 2010;375(9715):621-2.

94. Eaton JW, Hallett TB, Garnett GP. Concurrent sexual partnerships and primary HIV infection: a critical interaction. AIDS Behav. 2011;15(4):687-92.

95. Knopf A, Morris M. Lack of association between concurrency and HIV infection: an artifact of study design. J Acquir Immune Defic Syndr. 2012;60(1):e20-1; author reply e1.

96. Kretzschmar M, Morris M. Measures of concurrency in networks and the spread of infectious disease. Math Biosci. 1996;133(2):165-95.

97. Maughan-Brown B, Kenyon C, Lurie MN. Partner age differences and concurrency in South Africa: implications for HIV-infection risk among young women. AIDS and Behavior. 2014.

98. McCreesh N, O'Brien K, Nsubuga RN, et al. Exploring the potential impact of a reduction in partnership concurrency on HIV incidence in rural Uganda: a modeling study. Sex Transm Dis. 2012;39(6):407-13.

99. Morris M, Epstein H. Role of concurrency in generalised HIV epidemics. Lancet. 2011;378(9806):1843-4; author reply 5-6.

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

101. 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.

102. Bellan SE, Dushoff J, Galvani AP, Meyers LA. Reassessment of HIV-1 acute phase infectivity: accounting for heterogeneity and study design with simulated cohorts. PLoS Med. 2015;12(3):e1001801.

103. Pilcher CD, Tien HC, Eron JJ, Jr., et al. Brief but efficient: acute HIV infection and the sexual transmission of HIV. J Infect Dis. 2004;189(10):1785-92.

104. Gaydosh L, Reniers G, Helleringer S. Partnership concurrency and coital frequency. AIDS Behav. 2013;17(7):2376-86.

105. Sawers L. Measuring and modelling concurrency. J Int AIDS Soc. 2013;16:17431.

106. Sawers L, Isaac AG, Stillwaggon E. HIV and concurrent sexual partnerships: modelling the role of coital dilution. J Int AIDS Soc. 2011;14:44.

107. Jenness SM, Biney AA, Ampofo WK, Nii-Amoo Dodoo F, Cassels S. Minimal coital dilution in Accra, Ghana. J Acquir Immune Defic Syndr. 2015;69(1):85-91.

108. UNAIDS Reference Group on Estimates, Modelling and Projections. Consultation on Concurrent Sexual Partnerships: Recommendations. 2009. Available from: http://www.epidem.org/sites/default/files/content/resources/attachments/Concurrency meeting recommendations_Final.pdf. Accessed 5 Aug 2013.

109. Helleringer S, Mkandawire 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.

110. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc R Soc A. 1927;115:700-21.

111. Busenberg S, Castillo-Chavez C. A general solution of the problem of mixing of subpopulations and its application to risk- and age-structured epidemic models for the spread of AIDS. IMA J Math Appl Med Biol. 1991;8(1):1-29.

112. Keeling MJ, Eames KT. Networks and epidemic models. J R Soc Interface. 2005;2(4):295-307.

113. Kretzschmar M, Dietz K. The effect of pair formation and variable infectivity on the spread of an infection without recovery. Math Biosci. 1998;148(1):83-113.

114. Powers KA, Ghani AC, Miller WC, et al. The role of acute and early HIV infection in the spread of HIV and implications for transmission prevention strategies in Lilongwe, Malawi: a modelling study. Lancet. 2011;378(9787):256-68.

115. Goodreau SM, Carnegie NB, Vittinghoff E, et al. Can male circumcision have an impact on the HIV epidemic in men who have sex with men? PloS one. 2014;9(7).

116. Khanna AS, Roberts ST, Cassels S, et al. Estimating PMTCT's Impact on Heterosexual HIV Transmission: A Mathematical Modeling Analysis. PloS one. 2015;10(8).

117. Jenness SM, Goodreau SM, Morris M, Cassels S. Effectiveness of combination packages for HIV-1 prevention in sub-Saharan Africa depends on partnership network structure: a mathematical modelling study. Sex Transm Infect. 2016;92(8):619-24.

118. Roberts ST, Khanna AS, Barnabas RV, et al. Estimating the impact of universal antiretroviral therapy for HIV serodiscordant couples through home HIV testing: insights from mathematical models. J Int AIDS Soc. 2016;19(1):20864.

119. Hunter DR, Handcock MS, Butts CT, Goodreau SM, Morris M. ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. J Stat Softw. 2008;24(3):nihpa54860.

120. Jenness S, Goodreau SM, Morris M, Beylerian E. Package 'EpiModel'. Package 'EpiModel'. 2015.

121. Krivitsky MPN. Package 'statnet'. Package 'statnet'. 2014.

122. Eaton JW, Bacaër N, Bershteyn A, et al. Assessment of epidemic projections using recent HIV survey data in South Africa: a validation analysis of ten mathematical models of HIV epidemiology in the antiretroviral therapy era. The Lancet Global health. 2015;3(10):608.

123. Bacaër N, Pretorius C, Auvert B. An Age-Structured Model for the Potential Impact of Generalized Access to Antiretrovirals on the South African HIV Epidemic. Bulletin of Mathematical Biology. 2010;72(8):2180-98.

124. Klein DJ, Bershteyn A, Eckhoff PA. Dropout and re-enrollment: implications for epidemiological projections of treatment programs. Aids. 2014;28.

125. Johnson L, Dorrington R, Bradshaw D. Sexual behaviour patterns in South Africa and their association with the spread of HIV: insights from a mathematical model. Demographic Research. 2009;21(11):289-340.

126. Phillips AN, Pillay D, Garnett G, et al. Effect on transmission of HIV-1 resistance of timing of implementation of viral load monitoring to determine switches from first to second-line antiretroviral regimens in resource-limited settings. Aids. 2011;25(6):843-50.

127. Cambiano V, Bertagnolio S, Jordan MR. Transmission of drug resistant HIV and its potential impact on mortality and treatment outcomes in resource-limited settings. J Infect Dis. 2013;207(Suppl 2):S57-62.

128. Ott MQ, Bärnighausen T, Tanser F, Lurie MN, Newell M-LL. Age-gaps in sexual partnerships: seeing beyond 'sugar daddies'. AIDS (London, England). 2011;25(6):861-3.

129. Department of Health Republic of South Africa. South Africa Demographic and Health Survey 1998: Full Report. 1998. Available from: http://dhsprogram.com/pubs/pdf/FR131/FR131.pdf. Accessed 12 March 2016.

130. Hontelez JA, Lurie MN, Bärnighausen T, et al. Elimination of HIV in South Africa through expanded access to antiretroviral therapy: a model comparison study. PLoS medicine. 2013;10(10).

2. Antiretroviral treatment and sexual risk behavior in South Africa[1]

Kathryn Risher, Thomas Rehle, Leickness Simbayi, Olive Shisana, David D. Celentano

2.1. Abstract

The sexual behavior of individuals living with HIV determines the onward transmission of HIV. With the understanding that antiretroviral therapy (ART) prevents transmission of HIV, the sexual behaviors of the individuals not on ART with unsuppressed viral loads becomes of the greatest importance in elucidating transmission. We assessed the association between being on ART and sexual risk behavior among those living with HIV in a nationally representative population-based cross-sectional survey of households in South Africa that was conducted in 2012. Of 2,237 adults (aged 15-49) who tested HIV-seropositive, 667 (29.8%) had detectable antiretroviral drugs in their blood specimens. Among males, 77.7% of those on ART reported having had sex in the past year contrasted with 88.4% of those not on ART (p=0.001); among females, 72.2% of those on ART reported having had sex in the past year while 80.3% of those not on ART did (p<0.001). For males and females, the odds of reporting consistent condom use and condom use at last sex were statistically significantly higher for individuals on ART compared to those not on ART (Males: consistent condom use aOR=2.8, 95% CI=1.6-4.9, condom use at last sex aOR=2.6, 95% CI=1.5-4.6; Females: consistent condom use

1. Published as: Risher K, Rehle T, Simbayi L, Shisana O, Celentano DD. Antiretroviral treatment and sexual risk behavior in South Africa. *AIDS Behav*. 2016; 20(4):710-6.

aOR=2.3, 95% CI=1.7-3.1, condom use at last sex aOR=2.3, 95% CI=1.7-3.1), while there were no statistically significant differences in odds of reporting multiple sexual partners in the past year. In this nationally representative population-based survey of South African adults, we found evidence of less risky sexual risk behavior among people living with HIV on ART compared to those not on ART.

2.2. Introduction

South Africa has been greatly impacted by the HIV epidemic, with an estimated HIV prevalence among individuals aged 15-49 of 18.8% (95% confidence interval [CI]: 17.5-20.3%) in 2012[1]. An estimated 6.4 million people were living with HIV infection in South Africa in 2012[1], making South Africa the country with the largest absolute number of people living with HIV worldwide and nearly one-sixth of the burden worldwide[2]. The government of South Africa has made a strong commitment to providing HIV counseling and testing (HCT) to all individuals, and provided HCT to approximately 9 million South Africans between April 2012 and March 2013[3]. Additionally, the South African government has greatly expanded access to antiretroviral treatment[4] including increasing coverage of antiretroviral therapy (ART) to an estimated over 2 million people in mid-2012[1]. While South Africa has made great strides to decrease the burden of HIV, there remains significant stigma around HIV and HIV remains a major public health problem.

The HIV Prevention Trials Network (HPTN) 052 study clearly demonstrated the protective nature of ART on risk of HIV transmission, showing that HIV transmission between serodiscordant couples is reduced by 96% with early treatment of the seropositive partner[5]. The protective nature of ART on HIV transmission makes the sexual behaviors of individuals who are not on ART, who mostly have unsuppressed viral loads and are at increased risk of transmitting their virus, of the utmost importance to understand in order to implement effective positive prevention strategies.

The sexual behavior of individuals on ART is additionally of interest. People living with HIV who initiate ART often experience an improvement in general health and

may have an accompanied increase in sexual activity[6]. Given that superinfection with novel strains, particularly with drug-resistant strains of HIV, can speed up disease progression and complicate the treatment of HIV[7], understanding the risk behavior of people both on and off ART is necessary, particularly in the context of increasingly earlier treatment guidelines[8]. Increased sexual risk taking among individuals both on and not on ART, mainly increases in unprotected sex acts and increased number of sexual partners, puts individuals at increased risk for sexually transmitted infections (STIs) and unplanned pregnancies. Additionally, adherence to ART becomes increasingly important, including as protection from superinfection, though previously adherence has been found to be inconsistent in South African populations[9,10].

While early studies from Western countries suggested that individuals on ART may have increased risk behavior[11-14], most literature published in South Africa[15-18] and other settings in sub-Saharan Africa[19-27] (including a meta-analysis published in 2012[28]) found either no significant difference in sexual risk behavior or a decrease in sexual risk behavior between individuals on ART compared to those who are not, while a few studies showed support for an increase in sexual risk behavior by individuals on ART[29-31]. This question has yet to be assessed on a nationally representative scale in South Africa.

We aimed to assess the association between being on ART and sexual risk behavior among people of reproductive age (age 15-49) living with HIV in a nationally representative sample from South Africa undertaken in 2012. We assessed the following among those who reported sexual activity in the past 12 months: consistency of condom

use, condom use at last sex and multiple sexual partnerships, comparing those on and not on ART.

2.3. Methods

2.3.1. *Study Design and Setting*

The fourth South African National HIV Prevalence, Incidence, and Behaviour Survey (SABSSM IV) was a cross-sectional survey conducted in 2012 among individuals of all ages in all provinces of South Africa. SABSSM IV was a multi-stage cluster sample survey of residential households within enumeration areas sampled by province, race group and locality type (urban/rural and formal/informal) as defined by the census in South Africa[32]. Institutionalized individuals (including individuals in educational institutions, military barracks, old-age homes, or hospitals) were excluded from the study. This analysis is restricted to individuals aged 15-49 living with HIV.

Interviewers administered questionnaires to participants that included demographics, sexual history (including partner history, consistency of condom use and condom use at last sex), HIV testing history and risk perception, history of sexually transmitted infections, male circumcision, contraceptive use, and alcohol use, amongst others. All youth and adults who participated provided either written or verbal consent, including parent/guardian informed consent for youth under 18 years of age and youth verbal assent to have a blood specimen taken.

2.3.2. Laboratory Measures

Dried blood spot (DBS) specimens were collected from each participant who assented or consented using finger prick. Samples were tested for HIV using an enzyme

immunoassay (EIA) (Vironostika HIV Uni-Form II plus O, Biomeriux, Boxtel, The Netherlands), and samples that tested positive were retested using a second EIA (Advia Centaur XP, Siemens Medical Solutions Diagnostics, Tarrytown, NJ, USA). Any samples with discordant results on the first two EIAs were tested with a third EIA (Roche Elecys 2010 HIV Combi, Roche Diagnostics, Mannheim, Germany). Samples that tested positive for HIV-1 antibodies were tested for the presence of antiretroviral drugs (ARVs) using high performance liquid chromatography (HPLC) coupled with tandem mass spectrometry. Zidovudine, Nevirapine, Efavirenz, Lopinavir, Atazanavir, and Darunavir were detected using an Applied Biosystems API 4000 tandem mass spectrometer. The limit of detection was set to 0.2 micrograms/ml.

2.3.3. Statistical Analysis

We report socio-demographics of participants living with HIV by ART status stratified by sex, and test for differences between groups by ART status using the Pearson chi-square test of independence, in STATA 12 (Stata Corp., College Station, TX). To assess the association between ART status and sexual behavior outcomes (consistency of condom use with most recent partner, condom use at last sex, and multiple sexual partnerships), again stratified by sex, we used simple and multiple logistic regression for each outcome variable. Variables were included in multiple logistic regression analysis based on our literature review, and our final models adjusted for age category, race, income in the past month, locality type, partner type, education and alcohol use, as well as disclosure of HIV status to partner and alternative contraceptive method for condom use outcomes. Consistency of condom use was dichotomized to indicate "always" compared to "less than always" use.

2.4. Results

Of 2,237 adults between the ages of 15 and 49 who tested HIV-seropositive, 667 (29.8%) had detectable ART in their blood specimens. Table ***2.1*** shows demographic characteristics of respondents living with HIV on and not on ART, stratified by gender. Among both males and females, the majority of the individuals on ART were older than those who were not on ART (median age 38 and 35 vs 32 and 30 for males and females, respectively). Fewer participants on ART had a regular sexual partner than those not on ART (among males: 69.0% of those on ART and 79.8% of those not on ART; among females: 66.6% of those on ART and 74.0% of those not on ART). A slightly higher proportion of males and females on ART had received any income in the past month compared to those not on ART (69.7% vs 61.4% among males, 59.0% vs 54.6% among females). The distribution of males on ART was different by locality type (more in rural informal areas were on ART: 32.1% vs 22.6%; and less often men in rural formal areas were on ART: 12.8% vs 18.9%), while among females, no locality differences were seen by ART.

Table ***2.2*** shows HIV risk behavior reported by participants living with HIV by ART status, stratified by sex. Individuals on ART were less likely to report having been sexually active in the past year than those who were not on ART (77.7% compared to 88.4% among males, and 72.2% compared to 80.3% among females). Higher proportions of respondents on ART reported that they used condoms consistently and at last sex compared to those not on ART (consistent condom use: 57.4% vs 34.9% among males, and 53.7% vs 33.5% among females; condom use at last sex: 73.2% vs 45.3% among males, and 65.5% vs 44.5% among females). The proportion of respondents

reporting multiple partnerships in the past 12 months was slightly lower among those on ART compared to those not on ART (14.9% vs 20.9% among males; 5.1% vs 7.7% among females). Those respondents on ART reported less binge drinking (28.8% vs 34.1% among males; 8.0% vs 11.1% among females), and more respondents on ART reported no alcohol use in the past 12 months (58.8% vs 46.7% among males; 87.0% vs 79.4% among females). A greater proportion of those on ART disclosed their status to their sexual partner (62.3% vs 46.3% among males and 60.3% vs 53.9% among females). Of those who were not on ART, 52.1% of women and 39.6% of men had received an HIV test and their results in the past 12 months. Significantly more individuals on ART reported healthcare services use in the past 6 months compared to those not on ART (71.0% and 71.5% compared to 41.0% and 50.2%, for males and females, respectively).

Figure ***2.1*** shows the unadjusted and adjusted associations between ART and sexual risk behaviors. For males and females, the odds of reporting condom use were statistically significantly higher for individuals on ART compared to those not on ART, even after adjusting for potential confounders (Males: consistent condom use aOR=2.8, 95% CI=1.6-4.9, condom use at last sex aOR=2.6, 95% CI=1.5-4.6; Females: consistent condom use aOR=2.3, 95% CI=1.7-3.1, condom use at last sex aOR=2.3, 95% CI=1.7-3.1). There was no statistically significant difference in reporting of multiple sexual partners in the past year comparing those on ART to those not on ART after adjusting for other potential confounders for either males or females (among males: aOR=0.81, 95% CI=0.41-1.6; among females: aOR=0.94, 95% CI=0.52-1.7). We assessed for mediation and effect modification by exposure to healthcare services in the past year and did not find a large impact (<10%) as a mediator nor statistically significant interaction term. We

additionally stratified our analyses by respondents' having disclosed their HIV status to partners, and respondents' awareness of their partners' status, but found our conclusions did not change.

2.5. Discussion

In this nationally representative population-based survey of South African adults, we found evidence of increased sexual risk behavior among people living with HIV not on ART compared to those who are on ART. Individuals on ART were more likely to report condom use and disclosing their status to partners, and less likely to report hazardous alcohol use.

The above results indicate the necessity for HIV prevention interventions to target individuals who are not yet on ART with behavioral HIV prevention interventions. Individuals not on ART reported high levels of sexual risk. Nearly two-thirds of HIV-positive individuals not on ART reported inconsistent condom use and over half of the same individuals did not use a condom at last sex. This suggests that the individuals with the greatest potential for onward transmission of HIV, those who are not on ART who likely have higher viral loads, are engaging in behaviors that put their sexual partners at high risk of infection. However, identifying this population for behavioral prevention interventions is a major challenge, given the possibility that these individuals are simply less likely or willing to engage with public health or clinical medicine and the combined challenges of sexual behavior change while simultaneously linking a newly diagnosed positive individual to clinical care.

Individuals on ART also reported much higher levels of recent interaction with healthcare providers. It is possible that the lower reported levels of sexual risk behavior

can be explained by this interaction, though effect modification and mediation by having seen a healthcare provider in the past 6 months was not significant. South African ART counseling guidelines include risk reduction/prevention for positives[4]. It is unclear if the effect we observe is simply a case of individuals who interact with the healthcare system knowing the "right" answers and therefore our results potentially could be due to social desirability bias, or if the observed effect is a truly protective effect.

There are numerous potential limitations to the results presented. The cross-sectional nature of the data collected in SABSSM IV means that no assessments of temporality may be made for the associations examined. We did not ask participants if they were aware of their HIV status, and therefore HIV-1 antibody-positive participants not on antiretroviral therapy will represent a mixture of participants who are aware of their status but not yet on ART (potentially because their CD4+ cell count is above 350 cells/mm^3), aware of their status and non-adherent to their ART regimen, and individuals who are unaware of their status. We are unable to assess if differences in reported sexual behavior could be attributed to simply awareness of HIV status, or to being on ART, which is potentially a major confounder of our results. Previous studies have found that self-reported HIV status and self-reported ART use are often misreported when compared to biological HIV testing and ART detection[33,34], making assessing awareness of HIV status and HIV care in household-based cross-sectional studies very challenging. Individuals living with HIV but unaware of their status have been shown to engage in the highest risk behavior[35-37], and may thus increase the risk profile of the not on ART group in our study.

We are thus additionally unable to separate out the sexual behavior of participants who are non-adherent to ART. However, the individuals who are on ART but non-adherent to their medications may be of the greatest interest in assessing potential for "risk compensation," in which individuals engage in higher risk behaviors because they perceive themselves to be protected[38]. Future studies should differentiate between non-adherence and individuals not yet on ART. Further, we did not measure CD4 levels to determine whether those not on ART would meet local treatment guidelines. Nor did we measure whether individuals were aware of the protective nature of ART on HIV transmission. As the results of HPTN 052 become increasingly widespread, it will be vital to understand if risk compensation attenuates the effect of treatment as prevention[39].

Additionally, sexual behavior was ascertained through participant self-report. Given the wide-scale general population knowledge about HIV and the numerous health education campaigns that have been conducted in South Africa, there is substantial reason to expect that self-report may be more conservative than true sexual behavior (though there is some question as to whether females under-report sexual activity while males over-report sexual activity[40]). Of greatest concern is the potential that participants who are on ART and therefore receiving counseling on safe sex are more likely to under-report their sexual activity than those who are not in regular care. There is little guidance to correct for this potential differential misclassification.

Cross-sectional studies have several additional limitations in answering the question of differences in sexual behavior between those on and not on ART. Particularly, the individuals who initiate and maintain ART likely are inherently different

from the individuals who do not initiate ART or those who never get tested, and this unmeasured confounding may lead to un-interpretable results. As a result, future studies assessing differential sexual risk taking by ART status should have more rigorous longitudinal designs and assess the nuances in the population not on ART.

As "treatment as prevention" strategies are increasingly implemented, it will remain vital to understand the sexual behavior of individuals living with HIV whether on or not on ART. Implementing HIV prevention interventions to decrease sexual risk behaviors of individuals not on ART are of the greatest importance for curtailing the HIV epidemic, though to identify these individuals they will need to be tested, setting them onto the "continuum of care." Additionally, as the public becomes increasingly aware of the protective nature of ART against transmission, assessing the potential for increased risk behaviors among those on ART will also be of interest. Continued monitoring of the sexual behaviors of individuals living with HIV and implementation of behavioral strategies to modify risky sexual behaviors among these individuals remain key strategic targets for the HIV response.

2.6. References

1. Shisana O, Rehle T, Simbayi L, et al. South African National HIV Prevalence, Incidence and Behaviour Survey, 2012. Cape Town: HSRC Press; 2014.

2. Abdool Karim SS, Abdool Karim Q, Gouws E, Baxter C. Global epidemiology of HIV-AIDS. Infect Dis Clin North Am. 2007 Mar; 21(1):1-17.

3. South African National AIDS Council. Republic of South Africa Global AIDS Response Progress Report 2013. Available at: http://sanac.org.za/publications/reports/doc_download/61-ungass-south-africa-report. Accessed 2 Aug 2014.

4. Department of Health Republic of South Africa. South African Antiretroviral Treatment Guidelines 2013. Available at: http://www.sahivsoc.org/upload/documents/2013 ART Treatment Guidelines Final 25 March 2013 corrected.pdf. Accessed 2 Aug 2014.

5. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med. 2011 Aug 11; 365(6):493-505.

6. Wamoyi J, Mbonye M, Seeley J, Birungi J, Jaffar S. Changes in sexual desires and behaviours of people living with HIV after initiation of ART: implications for HIV prevention and health promotion. BMC Public Health. 2011; 11:633.

7. Redd AD, Quinn TC, Tobian AA. Frequency and implications of HIV superinfection. Lancet Infect Dis. 2013 Jul; 13(7):622-8.

8. World Health Organization HIV/AIDS Programme. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach 2013. Available at:

http://apps.who.int/iris/bitstream/10665/85321/1/9789241505727_eng.pdf. Accessed 1 Aug 2013.

9. Peltzer K, Friend-du Preez N, Ramlagan S, Anderson J. Antiretroviral treatment adherence among HIV patients in KwaZulu-Natal, South Africa. BMC Public Health. 2010; 10:111.

10. Maqutu D, Zewotir T, North D, Naidoo K, Grobler A. Determinants of optimal adherence over time to antiretroviral therapy amongst HIV positive adults in South Africa: a longitudinal study. AIDS Behav. 2011 Oct; 15(7):1465-74.

11. Stolte IG, Dukers NH, Geskus RB, Coutinho RA, de Wit JB. Homosexual men change to risky sex when perceiving less threat of HIV/AIDS since availability of highly active antiretroviral therapy: a longitudinal study. AIDS. 2004 Jan 23; 18(2):303-9.

12. Dukers NH, Goudsmit J, de Wit JB, Prins M, Weverling GJ, Coutinho RA. Sexual risk behaviour relates to the virological and immunological improvements during highly active antiretroviral therapy in HIV-1 infection. AIDS. 2001 Feb 16; 15(3):369-78.

13. Ostrow DE, Fox KJ, Chmiel JS, et al. Attitudes towards highly active antiretroviral therapy are associated with sexual risk taking among HIV-infected and uninfected homosexual men. AIDS. 2002 Mar 29; 16(5):775-80.

14. Tun W, Gange SJ, Vlahov D, Strathdee SA, Celentano DD. Increase in sexual risk behavior associated with immunologic response to highly active antiretroviral therapy among HIV-infected injection drug users. Clin Infect Dis. 2004 Apr 15; 38(8):1167-74.

15. Eisele TP, Mathews C, Chopra M, et al. High levels of risk behavior among people living with HIV Initiating and waiting to start antiretroviral therapy in Cape Town South Africa. AIDS Behav. 2008 Jul; 12(4):570-7.

16. Peltzer K, Ramlagan S. Safer sexual behaviours after 1 year of antiretroviral treatment in KwaZulu-Natal, South Africa: a prospective cohort study. Sex Health. 2010 Jun; 7(2):135-41.

17. Venkatesh KK, de Bruyn G, Lurie MN, et al. Decreased sexual risk behavior in the era of HAART among HIV-infected urban and rural South Africans attending primary care clinics. AIDS. 2010 Nov 13; 24(17):2687-96.

18. Eisele TP, Mathews C, Chopra M, et al. Changes in risk behavior among HIV-positive patients during their first year of antiretroviral therapy in Cape Town South Africa. AIDS Behav. 2009 Dec; 13(6):1097-105.

19. Apondi R, Bunnell R, Ekwaru JP, et al. Sexual behavior and HIV transmission risk of Ugandan adults taking antiretroviral therapy: 3 year follow-up. AIDS. 2011 Jun 19; 25(10):1317-27.

20. Dia A, Marcellin F, Bonono RC, et al. Prevalence of unsafe sex with one's steady partner either HIV-negative or of unknown HIV status and associated determinants in Cameroon (EVAL ANRS12-116 survey). Sex Transm Infect. 2010 Apr; 86(2):148-54.

21. Bateganya M, Colfax G, Shafer LA, et al. Antiretroviral therapy and sexual behavior: a comparative study between antiretroviral- naive and -experienced patients at an urban HIV/AIDS care and research center in Kampala, Uganda. AIDS Patient Care STDS. 2005 Nov; 19(11):760-8.

22. Moatti JP, Prudhomme J, Traore DC, et al. Access to antiretroviral treatment and sexual behaviours of HIV-infected patients aware of their serostatus in Cote d'Ivoire. AIDS. 2003 Jul; 17 Suppl 3:S69-77.

23. Luchters S, Sarna A, Geibel S, et al. Safer sexual behaviors after 12 months of antiretroviral treatment in Mombasa, Kenya: a prospective cohort. AIDS Patient Care STDS. 2008 Jul; 22(7):587-94.

24. Wandera B, Kamya MR, Castelnuovo B, et al. Sexual behaviors over a 3-year period among individuals with advanced HIV/AIDS receiving antiretroviral therapy in an urban HIV clinic in Kampala, Uganda. J Acquir Immune Defic Syndr. 2011 May 1; 57(1):62-8.

25. Thirumurthy H, Goldstein M, Zivin JG, Habyarimana J, Pop-Eleches C. Behavioral Responses of Patients in AIDS Treatment Programs: Sexual Behavior in Kenya. Forum Health Econ Policy. 2012 Apr 19; 15(2).

26. Yalew E, Zegeye DT, Meseret S. Patterns of condom use and associated factors among adult HIV positive clients in North Western Ethiopia: a comparative cross sectional study. BMC Public Health. 2012; 12:308.

27. Bunnell R, Ekwaru JP, Solberg P, et al. Changes in sexual behavior and risk of HIV transmission after antiretroviral therapy and prevention interventions in rural Uganda. AIDS. 2006 Jan 2; 20(1):85-92.

28. Berhan A, Berhan Y. Is the Sexual Behaviour of HIV Patients on Antiretroviral therapy safe or risky in Sub-Saharan Africa? Meta-Analysis and Meta-Regression. AIDS Res Ther. 2012; 9(1):14.

29. Olley BO. Higher-risk sexual behaviour among HIV patients receiving antiretroviral treatment in Ibadan, Nigeria. Afr J AIDS Res. 2008; 7(1):71-8.

30. Shafer LA, Nsubuga RN, White R, et al. Antiretroviral therapy and sexual behavior in Uganda: a cohort study. AIDS. 2011 Mar 13; 25(5):671-8.

31. Pearson CR, Cassels S, Kurth AE, Montoya P, Micek MA, Gloyd SS. Change in sexual activity 12 months after ART initiation among HIV-positive Mozambicans. AIDS Behav. 2011 May; 15(4):778-87.

32. Statistics South Africa. Census 2011 South Africa 2012. Available at: http://www.statssa.gov.za/Publications/P03014/P030142011.pdf. Accessed 2 Aug 2014.

33. Fishel J, Barrere B, Kishor S. Validity of data on self-reported HIV status and implications for measurement of ARV coverage in Malawi. DHS Working Papers No. 81. Calverton, Maryland, USA: ICF International; 2012.

34. Fogel JM, Wang L, Parsons TL, et al. Undisclosed antiretroviral drug use in a multinational clinical trial (HIV Prevention Trials Network 052). J Infect Dis. 2013 Nov 15; 208(10):1624-8.

35. Dokubo EK, Shiraishi RW, Young PW, et al. Awareness of HIV status, prevention knowledge and condom use among people living with HIV in Mozambique. PLoS One. 2014; 9(9):e106760.

36. Rosenberg NE, Pettifor AE, De Bruyn G, et al. HIV testing and counseling leads to immediate consistent condom use among South African stable HIV-discordant couples. J Acquir Immune Defic Syndr. 2013 Feb 1; 62(2):226-33.

37. Bunnell R, Opio A, Musinguzi J, et al. HIV transmission risk behavior among HIV-infected adults in Uganda: results of a nationally representative survey. AIDS. 2008 Mar 12; 22(5):617-24.

38. Cassell MM, Halperin DT, Shelton JD, Stanton D. Risk compensation: the Achilles' heel of innovations in HIV prevention? BMJ. 2006 Mar 11; 332(7541):605-7.

39. Venkatesh KK, Flanigan TP, Mayer KH. Is expanded HIV treatment preventing new infections? Impact of antiretroviral therapy on sexual risk behaviors in the developing world. AIDS. 2011 Oct 23; 25(16):1939-49.

40. Clark S, Kabiru C, Zulu E. Do men and women report their sexual partnerships differently? Evidence from Kisumu, Kenya. Int Perspect Sex Reprod Health. 2011 Dec; 37(4):181-90.

Table 2.1: Demographic characteristics by ART status stratified by sex among South Africans living with HIV, 2012

	Males				Females			
	Not on ART (N= 509)		On ART (N=156)		Not on ART (N=1061)		On ART (N=511)	
	N	%	N	%	N	%	N	%
Age								
Median	32		38		30		35	
15-24	84	16.5	9	5.8	252	23.8	54	10.6
25-34	236	46.4	47	30.1	460	43.4	183	35.8
35-49	189	37.1	100	64.1	349	32.9	274	53.6
Race								
African	453	89.0	137	87.8	982	92.6	474	92.8
Non-African	56	11.0	19	12.2	79	7.5	37	7.2
Partner Status[a]								
No regular partner (Single, divorced, etc)	102	20.2	48	31.0	273	26.0	170	33.4
Regular partner (Married, cohabitating, going steady)	403	79.8	107	69.0	776	74.0	339	66.6
Received any income in last month[a]	309	61.4	108	69.7	569	54.6	298	59.0
Locality Type								
Urban Formal	198	38.9	50	32.1	351	33.1	162	31.7
Urban Informal	100	19.7	36	23.1	218	20.6	112	21.9
Rural Informal	115	22.6	50	32.1	371	35.0	182	35.6
Rural Formal	96	18.9	20	12.8	121	11.4	55	10.8

[a] Totals within demographic characteristics do not add to overall total due to missing data

Table 2.2: HIV risk behaviors by ART status stratified by sex among South Africans living with HIV, 2012

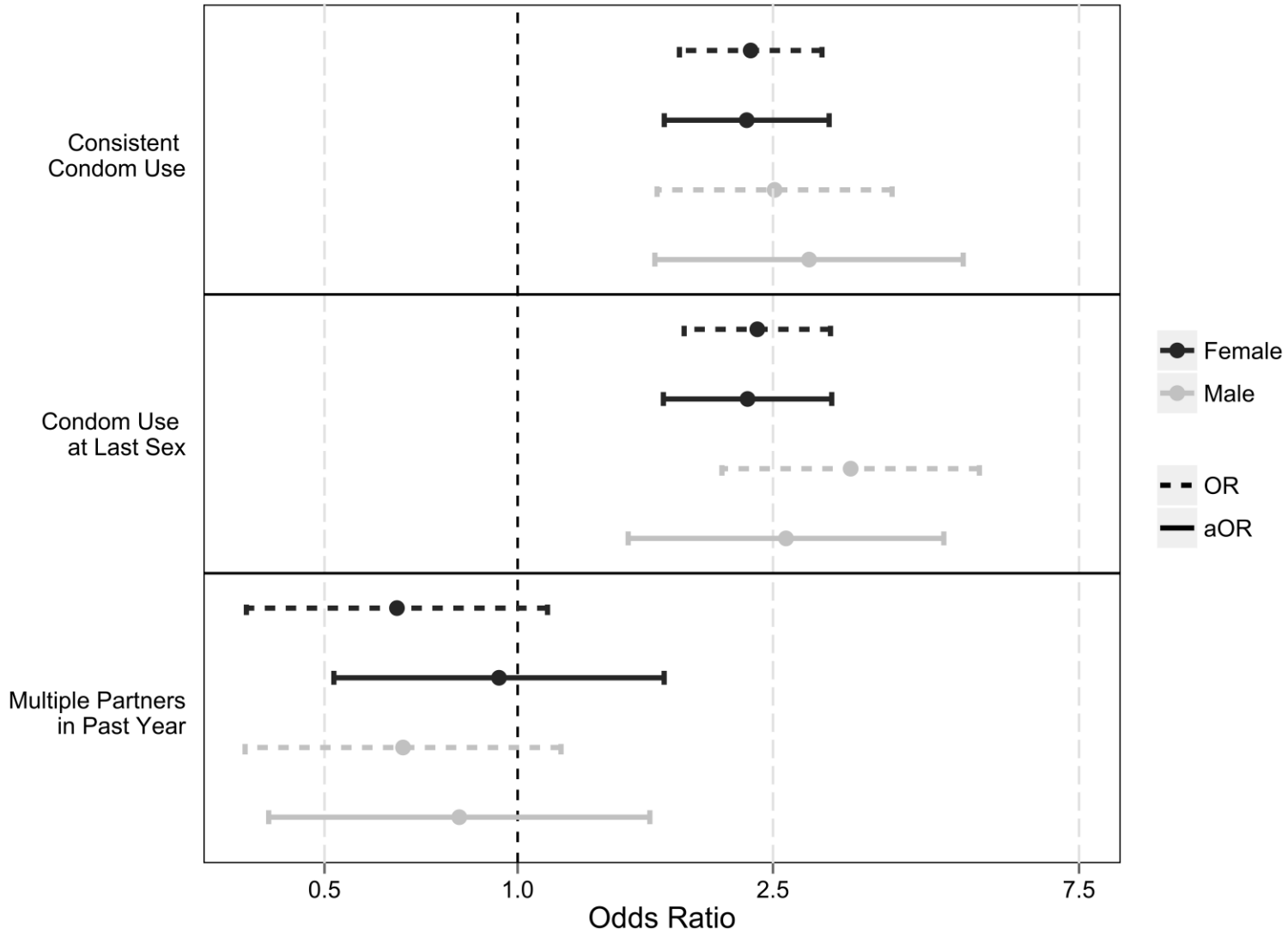
	Males					Females				
	Not on ART (N= 509)		On ART (N=156)		Chi-squared statistic (df)	Not on ART (N=1061)		On ART (N=511)		Chi-squared statistic (df)
	N	%	N	%	[p-value]	N	%	N	%	[p-value]
Sex in past 12 months	417	88.4	115	77.7	10.5 (1) [0.001]	813	80.3	351	72.2	12.2(1) [<0.001]
Most recent sex partner type[a]					5.8 (2)					0.8 (2)
Husband/Wife/Live-in Partner	169	41.3	61	53.0	[0.054]	306	38.1	141	40.2	[0.673]
Girlfriend/Boyfriend not live-in	233	57.0	51	44.4		479	59.6	200	57.0	
Consistency of Condom Use[a]					31.7 (3)					51.5 (3)
Every time	143	34.9	66	57.4	[0.001]	271	33.5	188	53.7	[<0.001]
Almost every time	22	5.4	12	10.4		41	5.1	18	5.1	
Sometimes	59	14.4	16	13.9		162	20.0	67	19.1	
Never	186	45.4	21	18.3		336	41.5	77	22.0	
Condom at last sex[a]	182	45.3	82	73.2	27.4 (1) [<0.001]	359	44.5	224	65.5	42.2 (1) [<0.001]
Multiple Partners[a]	87	20.9	17	14.9	2.04 (1) [0.153]	62	7.7	18	5.1	2.5 (1) [0.114]
Binge Drinking					7.6 (2)					14.3 (2)
No drinking in past 12 months	233	46.7	90	58.8	[0.023]	834	79.4	435	87.0	[0.001]
Drinking with no binge drinking	96	19.2	19	12.4		100	9.5	25	5.0	
Binge drinking	170	34.1	44	28.8		117	11.1	40	8.0	

Disclosed Status to all current partners					39.1 (2)					44.9 (2)
No	221	44.8	28	18.2	[<0.001]	301	29.1	71	14.3	[<0.001]
Yes	228	46.3	96	62.3		557	53.9	299	60.3	
No partner	44	8.9	30	19.5		175	16.9	126	25.4	
Received HIV test and results in past year	197	39.6	N/A	N/A	N/A	539	52.1	N/A	N/A	N/A
Exposure to Healthcare services in past 6 months	206	41.0	110	71.0	42.8 (1) [<0.001]	526	50.2	358	71.5	62.6 (1) [<0.001]

[a] Denominator represents only individuals who report being sexually active in the past 12 months

Figure 2.1: Unadjusted and adjusted odds ratios for sexual behaviors by ART status among South African males and females living with HIV, 2012

All aORs are adjusted for age category, race, income in past month, locality type, partner type, education, and alcohol use, while condom use outcomes additionally adjust for disclosure of HIV status to partner and alternative contraceptive use


Consistent
Condom Use
Condom Use
at Last Sex
Multiple Partners
in Past Year
0.5
1.0
2.5
7.5
Odds Ratio
Female
Male
OR
aOR

3. Sexual behavior and mixing patterns in South Africa, 2012

Kathryn A. Risher, Justin Lessler, Elizabeth Ogburn, Shruti H. Mehta, Olive Shisana, Thomas Rehle, Leickness Simbayi, Sizulu Moyo, Tarylee Reddy, David D. Celentano

3.1. Abstract

Background: Sexual transmission accounts for most HIV transmission in sub-Saharan Africa, but little is known about the structure of sexual networks, and the resulting transmission networks, in the region. We examined sexual mixing and sexual partnership distributions in South Africa.

Methods: Egocentric and household sexual relationship data came from a nationally representative household-based survey of South Africans conducted in 2012. We use mixing matrices, Newman's assortativity coefficients, and fit distributions to characterize like-with-like sexual mixing patterns and number of sexual partners.

Results: Among 15,283 sexually active respondents, the number of sexual partners reported in the last year was low (mean = 1.35), and men were 5.1 times as likely to report ≥2 partners than among women. A discrete q-exponential distribution best fit observed non-live-in number of partners. We found substantial assortativity by age, race, education, HIV-status, number of sexual partners and ARV use (assortativity coefficients>0.6).

Conclusions: Our empirical results can be used to understand HIV transmission in South Africa, including the parameterization of mathematical modeling studies of sexual mixing and HIV transmission to model network impacts on interventions. Better estimates of sexual mixing patterns and degree distributions will allow for more targeted HIV prevention efforts.

3.2. Introduction

In sub-Saharan Africa (SSA), HIV primarily is spread through sexual contact [1, 2], thus an understanding of the structure and dynamics of sexual networks aids the understanding and prevention of HIV transmission. There are limited data on sexual networks in the general heterosexual population of SSA. Data on sexual networks is primarily collected through egocentric methods, in which a respondent is asked about their network contacts but those contacts are not necessarily included in the study, or sociometric methods, in which all network members are included [3]. Sociometric data are more complete and allows for the direct examination of higher order network structure, but can be quite challenging to collect. Egocentric data, on the other hand, are easily collected. While they do not allow assessment of full network structure, they do allow an assessment of sexual mixing patterns and degree (number of contacts)[4].

With the exception of a detailed sociometric network study conducted on Likoma Island, Malawi [5-7], and a large body of work conducted on age-mixing between young women and older men [8-20], relatively few studies in SSA have focused on sexual mixing. Assortativity, partnering of like-with-like, can greatly impact rate of increase and ultimate burden of an HIV epidemic in a population, though the impact of assortativity varies by epidemic stage and mixing characteristic assessed[21, 22]. Further, an understanding of sexual mixing patterns can help target and implement interventions [23, 24]. This is of particular importance in South Africa, which has the highest absolute number of people living with HIV of any country in the world [25], and has been an extensive focus in the HIV modeling community.

In recent mathematical models of HIV transmission in South Africa [26-38], some models incorporate age assortativity [27-31, 38] from a variety of sources [39-41]. Other models included sexual activity level mixing, i.e., between low- and high-level sexual risk groups, though these parameters were exclusively calibrated, as opposed to empirically derived [32-35]. None of these models have incorporated sexual mixing on characteristics other than age and sexual risk level.

Sexual mixing in numerous settings is often characterized using Newman's assortativity coefficient [42-55]. This measure was recently used to assess mixing by educational attainment between partners using two rounds of Demographic and Health Surveys (DHS) data from Cameroon, Ethiopia, Kenya, Lesotho, Malawi, Rwanda, and Zimbabwe, where assortativity coefficients for education ranged from 0.09 in urban Zimbabwe to 0.44 in non-urban Cameroon, on a scale from (at lowest) -1 to 1 where values over 0 indicate assortativity [56]. Using other measures, a South African study assessed sexual mixing by ethnicity in the 1999 Carletonville Youth Survey [57], finding that ethnic mixing of sexual partners was homophilous (homophily and assortativity are used interchangeably) among 15-24 year old participants, though the extent to which this was observed varied by ethnicity.

Research in SSA has generally shown that people report a low number of lifetime and past year sexual partnerships. In an assessment of nationally representative surveys of adults from Kenya, Lesotho, Uganda and Zambia, Morris et al. found that the average proportion reporting two or more partners in the past year was 11.4 to 19.8% (depending on study administrator – PSI vs DHS), but that the proportion was dramatically lower among women (2.8 to 9.2%) compared to men (21.5 to 32.3%) [58].

Sexual degree data frequently have a strong right tail[59], with a high proportion of individuals having zero or one partner and a minority having much higher numbers of partners. Authors have shown that R_0, the number of new infections resulting from a single infected individual in a fully susceptible population, increases with the variance of number of contacts (in HIV, number of sexual partners) divided by the mean number of contacts [60-62]. An influential paper suggested that past year sexual partnerships in Sweden follow a scale-free power law distribution, with infinite variance, consistent with a preferential-attachment mechanism, and suggestive that there was no epidemic threshold for sexually transmitted diseases in this population [63]. An epidemic threshold determines whether a disease is controllable. Additional studies have found that degree distributions in SSA also follow a similar scale-free or close to scale-free distribution [64, 65], though others have found that distributions with finite variance fit best [66]. However, other authors have found that some methods to fit power law distributions lead to inaccurate results, frequently power laws only apply to certain values of degree above a minimum, and it is rare to be able to state definitively that data arose from a power law distribution[67], suggesting a concrete epidemic threshold [66, 68]. Most recent work fitting distributions to empirical degree data has focused on finding the best count distributions to fit collected data [69-71], though mitigating transmission by intervening among those with the highest number of partners remains the public health strategy regardless of the distribution fit [72, 73].

In the context of limited data on sexual mixing in southern Africa, we analyze a dataset of egocentric and household-based sexual partnerships collected in a nationally representative survey of South Africans in 2012. We describe reported sexual behavior

and present network statistics including mixing matrices, assortativity coefficients and degree distributions.

3.3. Methods

3.3.1. South African National HIV Prevalence, Incidence and Behaviour Survey Description

The South African National HIV Prevalence, Incidence and Behaviour Survey (SABSSM IV) was conducted in 2012. The survey was a multi-stage cluster sample of households in South Africa, and the survey methodology has been previously described [25]. Briefly, 1,000 enumeration areas were selected sampled by province, locality type (urban formal, urban informal, rural formal and rural informal), and race strata[2]. Within each enumeration area, 15 households were selected to participate using aerial photos. Within each selected household all individuals were invited to participate. Up to five visits were made at each household to ensure full participation. Dried blood spot specimens (DBS) were collected from every participant who consented and tested for HIV using serology, and the DBS that tested HIV-positive were tested for the presence of ART.

Respondents aged 15 and older were asked the number of sexual partners they have had in their lifetime and in the past year. For their most recent sexual partners in the past year (up to three in the past year), participants were asked to list the age and type of partner (husband/wife, live-in partner, girlfriend/boyfriend not living with you, casual

[2] Although legally mandated racial discrimination was abolished in 1994, the four racial categories of Black African, White, Coloured and Indian/Asian continue to be used as both social and economic inequalities still prevail as a legacy of the apartheid policies.

partner, someone whom you paid for sex, other), the start and end dates of the partnership, and patterns of condom use (at last sex and consistency).

Respondents with household sexual partners who also completed the SABSSM IV questionnaire were asked to identify this partner at the end of the survey, and the partner's study id number was recorded. Using this variable, an entire participant's questionnaire could be tied to his or her partner's questionnaire. In matching household partners, we did not require that both partners have listed one another but included partnerships where only one partner listed the other.

Our results are based on three different sources within SABSSM IV: self-reported number of partners (degree distribution), self-reported egocentric/partnership history data on age and partner type of three most recent sexual partners in past year, and household partners' linked questionnaires. Age assortativity for live-in partners is presented twice, once from household results and once from egocentric data. The reasons for presenting this from both data sources are multi-fold. First, household linked questionnaires represent the participants' true ages, as opposed to egocentric data where the alter's (partner's) age is reported by the ego (respondent). Second, the egocentric data on live-in partners is more extensive than the household partnership data. Respondents could report on live-in partners in the egocentric partnership history even if their partner was, for instance, a migratory worker and thus away during the time of the survey, or their partner refused to participate, or both partners chose not to disclose their household sexual partner at the end of the questionnaire to allow for linkage. Third, while household-linked partners are represented in the egocentric live-in partner data who we would be able to identify through the linked questionnaires as one another's partners through assessing

concordance between linked questionnaire partner age and egocentric partner age (see further analysis in appendix), we are unable to decisively conclude that a live-in partner is, in fact, the linked partner due to potential misreporting. Therefore, while the live-in partners reported in a questionnaire that was linked to another household member most likely represents that same live-in partnership, there will be limited cases where it represents an additional live-in partner. Further, due to an inability to tease apart egocentric live-in partners that are both included in the survey (male and female reports that are not independent data points) and egocentric live-in partners that are not both included in the survey, we analyze egocentric age mixing data separately by sex.

We restrict our analyses of number of partners to individuals aged 15-65, because individuals over 65 are substantially less likely to report sexual activity in the past year (among 2,239 participants aged >65, 76.7% with non-missing data report no sexual partners during the past year). For mixing analyses, egocentric and household partnership data are presented for all ages, instead of for a subset of ages, to avoid artificial edge effects created by excluding partners over 65, under the assumption that partnership mixing of individuals over age 65 are not systematically different from those under the age 65.

All youth and adults who participated provided either written or verbal consent, including parent/guardian informed consent for respondents less than 18 years of age. SABSSM IV has Institutional Review Board (IRB) approval from the Human Sciences Research Council IRB and the Centers for Disease Control and Prevention Center for Global Health IRB.

3.3.2. Data analysis

Throughout our analyses, we present the egocentric partner census separately from household partnerships. We refer to "household" and "non-household" partners with "household" representing marital or live-in partnerships and "non-household" referring to all other types of partnerships (girlfriend/boyfriend, casual, paid, other). Data is weighted at the individual and household-level to account for the complex survey design and survey refusal. Throughout our analyses, we assume that observed data is representative of missing observations (see appendix for description of missingness). We assume that missing egocentric partners are missing completely at random. All analyses were conducted in R version 3.3.1 (R Foundation for Statistical Computing).

To estimate the relative increase in same group mixing relative to random mixing stratified by sex for categorical variables (and in analysis of continuous variables categorically), we divide the proportion of observed in-group mixing by the proportion of the opposite sex sample that is in that group. To estimated confidence intervals on this relative increase, we conducted 1,000 weighted bootstrapped samples of our data and recalculated the statistic for each sample. A very small number of household partners are in more than one household partnership, and thus we assume that choice of additional household partners is independent of initial household partner in order for our bootstrapped confidence intervals to be valid.

To estimate the count of "non-household" partnerships in the past year, we subtracted the number of live-in and marital partners designated in the egocentric data section (of 3 most recent partners) from the total number of partners in the past year.

This assumes that any live-in or marital partners would have been in the three most recent sexual partners for any respondent.

3.3.3. Mixing Matrices

To develop sexual mixing matrices from egocentric and household partnership data, we calculated the relative increase in observed number of partnerships in two groups over expected number of partnerships between people in the two groups based on the marginal number of partnerships in each group. Hence, each cell represents the increase in number of partnerships between given groups relative to random mixing [74, 75].

$$\frac{Observed}{Expected} = \frac{N_{ij}}{\frac{N_{i.}}{N} * \frac{N_{.j}}{N} * N} = \frac{N_{ij} * N}{N_{i.} * N_{.j}}$$

Where N_{ij} represents the number of partnerships observed between groups i and j, $N_{i.}$ represents the number of partnerships reported by people in group i (the row sum), and N represents all of the partnerships reported. Each cell's N_{ij} is weighted and as such our results represent a weighted pseudopopulation (described in appendix).

We use a Mantel test [76] to test for differences between two mixing matrices (described in appendix).

3.3.4. Assortativity Coefficients

We use two formulations of Newman's assortativity coefficient [77, 78], one for discrete characteristics (race, education, employment, HIV status, ART status, HIV testing history) and another for continuous characteristics (age, degree). The discrete assortativity coefficient is:

$$r = \frac{\mathrm{Tr}\,\mathbf{e} - \|\mathbf{e}^2\|}{1 - \|\mathbf{e}^2\|}$$

Where **e** is the adjacency matrix between characteristics divided by the sum of the matrix (i.e., each element e_{ij} is the fraction of all partnerships that connect a partner of characteristic *i* to one of characteristic *j*) and ||**x**|| denotes the sum of all elements in matrix **x**. We use Newman's closed form variance formula (an intraclass correlation variance estimate):

$$\sigma_r^2 = \frac{1}{M}\frac{\sum_i a_i b_i + \left[\sum_i a_i b_i\right]^2 - \sum_i {a_i}^2 b_i - \sum_i a_i {b_i}^2}{1 - \sum_i a_i b_i}$$

Where M represents the number of edges, and $\sum_j e_{ij} = a_i$ and $\sum_i e_{ij} = b_j$, the marginal sex-specific population proportions in a given group (e.g. proportion of men with primary education). We also present jackknife and weighted bootstrap confidence intervals.

For continuous variables, the assortativity coefficient is equal to:

$$r = \sum_{xy} \frac{xy\left(e_{xy} - a_x b_y\right)}{\sigma_a \sigma_b}$$

Where *x* and *y* represent the values of the continuous variables, e_{xy} is the value of the matrix **e** above between partners with characteristics *x* and *y*, a_x and b_y are the marginal proportion of partnerships among those with characteristics *x* and *y*, and σ_a and σ_b are the standard deviations of the distributions of a_x and b_y, with the variance denoted by:

$$\sigma_a^2 = \sum_x x^2 a_x - \left[\sum_x x a_x\right]^2$$

Newman's continuous coefficient is a Pearson correlation coefficient, and we use Fisher's transformation [79] which showed that the inverse hyperbolic tangent transformed correlation coefficient has approximate standard error $\frac{1}{\sqrt{N-3}}$ to calculate

confidence intervals. We also present jackknife and weighted bootstrap confidence intervals.

Newman's coefficients are bounded by -1 and 1 (though the categorical coefficient only approaches -1 for a variable with 2 categories [80]), with 0 representing random mixing and 1 representing perfectly assortative mixing. Newman's coefficients avoid the problems presented by Q-statistics/modularity (which are not bounded) by normalizing the quantity and scaling it to range from -1 to 1. Previous authors have interpreted Newman's assortativity coefficients of 0.35 or higher as highly assortative, 0.15 to 0.34 as assortative, 0.10 to 0.15 as minimally assortative and below 0.10 as disassortative [44, 48]. For egocentric data, we assume that individuals choose each partner independently of the age of their previous partners. To test this assumption, we compare assortativity coefficients calculated with only most recent sexual partner. We made the same assumption with household partners that report polygamy/polyamory, though we do not test this assumption due to it representing an extremely small proportion of the household partnership sample. For both assortativity coefficients, our confidence intervals do not represent our uncertainty in the sampling weights, but rather take weights as constants and present confidence around our weighted pseudopopulation estimate.

3.3.5. Degree Distribution

A number of count distributions have been proposed to fit sexual partnership degree distributions [63, 66, 69, 81, 82]. We fit distributions to our non-household and household partnership degree counts for sexually active individuals (1+ partners in the past year) separately for men and women. The rationale for choosing to fit to degree 1+ is

to allow comparability between distributions that allow 0's (such as Poisson and negative binomial) and distributions that disallow 0's (such as discrete Pareto and Yule). We fit the following distributions: Poisson distribution, Poisson with lognormally distributed mean, negative binomial distribution, geometric distribution, discrete Pareto distribution, Waring distribution, Yule distribution, discrete q-exponential, geometric-Waring distribution, geometric-Yule distribution, negative binomial-Waring distribution, and negative binomial-Yule distribution. These distributions' probability mass functions (PMFs) are listed in the appendix.

We modified the R package 'degreenet' to allow for weighted degree distributions [82, 83]. For distributions which allow 0's (Poisson, negative binomial), the package standardizes the remaining probability mass function (PMF) by the proportion greater than 0 (i.e., divides the remaining PMF by 1-P(K=0)), and does the reverse for distributions with an artificial upper bound. Our upper bound for the household partnerships was 3 because of the structure of the egocentric partnership census. We estimated the maximum likelihood parameters for each of these distributions given our data, then compared the AICc (AIC with correction for finite sample sizes)

$$AICc = 2k - 2\log(\mathrm{L}) + \frac{2k(k+1)}{n-k-1}$$

of the potential distributions to select the distribution that we utilized. Due to the highly skewed nature of the household partnership data (almost exclusively 1's), we were unable to fit all distributions to the household partnership data (those with the Yule distribution were unable to be fitted).

3.4. Results

Among 38,098 respondents in SABSSM IV, 24,367 were aged 15-65, and 15,283 (68.3%) reported at least one sexual partner in the previous year. The overall number of sexual partners reported was low (mean = 0.92, among sexually active mean = 1.4). Figure 3.1 shows the full degree distributions for number of sexual partners in the past year (Figure 3.1a) and number of sexual partners in the respondents' lifetime (Figure 3.1b) among all respondents stratified by sex. Figure 3.2a shows that the proportion of males reporting two or more partners in the past year is much higher than among females (5.1 times higher), and that the proportion reporting two or more partners in the past year decreases with increasing age. Figure 3.2b shows that number of lifetime partners reported is higher for men than women in all age groups. The average age of sexual debut among participants was 18.5 years of age (median 18), and Figure 3.2b shows that the proportion with 0 lifetime partners rapidly drops between the 15-20-year age group and the 21-25-year age group.

In Table S3.1, we present the maximum likelihood estimates for fitted distributions to the non-household and household degree data. The discrete q-exponential distribution best fit the non-household degree data for both men and women, while the geometric-Waring (women) and discrete Pareto (men) distributions best fit the household degree data. Figure S3.1 shows the fitted distributions and observed data for non-household (Figure S3.1a and b) and household (Figure S3.1c and d) partnerships stratified by sex. Very few individuals with at least one live-in partner reported more than one live-in partner (among women, 0.3%, among men, 1.0%).

3.4.1. Household results

We were able to match 5,217 couples based on partner unique IDs collected at the end of the survey. We found that 91 (1.7%) household same-sex partnerships were reported (56 female-female, 35 male-male). Due to the potential for information bias as a result of same-sex sexual behaviors being highly stigmatized and the differential impact of this stigma in different settings in South Africa, we anticipate that the same-sex partnerships captured in this sample is biased with little possibility to mitigate this bias. As such, we exclusively analyze the 5,126 (98.3%) heterosexual partnerships reported. Seven men were listed as a household partner by two women, an additional one man was listed as a household partner by three women. Four women were listed as a household partner by two men. Of all participants with appropriately specified household partners, 96.6% (8940/9253) reported their relationship status as married and living with their partner, living with their non-marital partner, or in a civil union (marital status was missing for 999 of the household partners). Of the additional 313 individuals, 81 (25.9%) reported being married but not living with their spouse, and 132 (42.2%) reported having a boyfriend/girlfriend with whom they did not reside.

Females in household partnerships ranged in age from 13 to 99 with a median of 44 (IQR=33-54) while males ranged from 15 to 103, median 47 (IQR=37-58). Male partners were on average 3.8 (median 3, IQR = 1-6, range = -58 to 56) years older than female partners. Figure 3.3a and b show the age difference between household partners by male and female age.

Table 3.1 shows a description of the household partnerships, the percent of household partnerships that are “in group” or between two individuals with the same

characteristic, and the relative increase in "in group" over what would be anticipated due to random mixing, stratified by sex. In the household sample, black African race made up the largest group (63.1%), with coloured, Indian/Asian and white representing from 4-20% each. Just over half of individuals had received some secondary education or completed secondary school. The proportions employed varied substantially between men and women in household partnerships, with 58.3% of men reporting employment compared to only 38.4% of women. One-tenth of the household sample was living with HIV. Nearly 40% of female household partners reported only one lifetime partner, while 20% of males reported only one lifetime partner. In the past year, 82 and 86% of male and female household partners had one partner, while only 5.4 and 1.7% of men and women reported 2 or more. The proportion of household partners in which both partners were in the same group (% in-group) was very high for race (93-98%), slightly lower for education (68-83%), and lower still for employment (50-80%). Fifty- to seventy-percent of HIV positive individuals were in a relationship with someone else living with HIV, and 92-96% of HIV-negative individuals were with another negative partner. In-group percent for lifetime number of sexual partners ranged from 12 to 81%, and for sexual partners in the past year ranged from 11 to 96%. The relative increase in in-group assortativity compared to random mixing was above one for all groups, though the magnitude of the increase varied substantially. The relative increase was highest overall for race (ranging from 1.6-24.1). Any relative increase with a lower confidence bound over 1 suggests that we observe significantly more within-group mixing than is expected by chance alone. However, the size of the relative impact is determined partially by the size of the group (i.e. a very small group can have a very large relative increase, while a

very large group cannot). Relative increase was also quite high for HIV-positive individuals (5.0). Notably, it was very high for having had 2+ sexual partners in the past year (5.9 and 6.4), suggesting evidence of assortativity by degree, though there was not particularly strong evidence of assortativity by lifetime degree.

Figure 3.4 shows the mixing matrices for the household partnership data, while Table 3.2 shows assortativity coefficients for age mixing in the household partnership data. In Figure 3.4a, age mixing is strongly assortative, with male partners slightly older than female partners consistently across ages, corresponding to an assortativity coefficient of 0.88. b shows strongly assortative mixing by race (assortativity coefficient, AC=0.96). Figure 3.4c shows less assortativity by education, with a great deal of mixing between those with a secondary education and those with primary or tertiary, but very little mixing between tertiary and primary educated individuals (AC=0.68). Figure 3.4d shows mixing by employment status; while mixing is greatest among those who are more similar in their employment status, slightly more employed males are partnered with unemployed females than there are employed females with unemployed males (AC=0.48). Figure 3.4e shows the mixing matrix by HIV status, with much higher than expected mixing between individuals living with HIV (AC=0.67). Figure 3.4f shows mixing by lifetime number of partners, and shows a higher proportion of mixing among those with 8+ lifetime partners than would have been expected (AC=0.28 for categorical, 0.32 for continuous). Figure 3.4g shows mixing by number of partners in the past year (AC=0.72 for categorical, 0.19 for continuous). The assortativity coefficient for categorical number of partners in the past year is substantially higher than the continuous estimate, and likely is more meaningful because a precise correlation between number of

partners is likely not as indicative of HIV network risk as categorization into two or more partners vs. one or zero. Figure 3.4h shows somewhat assortative mixing by HIV testing history (AC=0.43). Figure 3.4i shows that individuals on ART are more likely to be partnered with others on ART, if both are positive (AC=0.70). All assortativity coefficients were high (0.19-0.96) and all qualify in the range of "assortative" or "highly assortative" (described in Methods).

3.4.2. Egocentric results

Among 16,581 egocentric partnerships reported by respondents, 354 (2.1%) were between same-sex partners (146 female-female and 208 male-male). Eight (2.3%) of persons who reported same-sex partnerships also reported a heterosexual partnership in the past year. Once again, we exclusively analyze the sample of opposite sex partnerships (including the eight individuals who had both a same- and opposite-sex partner in the past year), leaving 16,227 opposite sex partnerships reported by 15,872 respondents. Of these 16,227 partnerships, 7,058 (44.5%) were with a marital partner, 1,863 (11.8%) were with a non-marital live-in partner, 6,170 (38.9%) were with a boyfriend/girlfriend with whom the partner did not live, 662 (4.2%) were with a casual partner, 10 (0.06%) were with an individual the respondent had paid for sex, and 88 (0.6%) categorized their partnership as "other." Males reported being an average of 3.7 (median = 3, IQR= 1-6, range= -58 to 65) years older than the female partner in egocentrically collected partnerships. This difference varied minimally by partner type; the average difference was 3.7 (median=3, IQR=1-6, range= -53 to 61) among live-in partners, 3.5 (median=3, IQR=1-5, range= -58 to 47) among non-live-in boyfriend/girlfriend, and 4.1 (median=3, IQR=1-6, range = -56 to 65) among casual

partnerships. Figure 3.3c-f shows the distribution of age differences by respondent age in the egocentric partnership data. While age difference appears to be fairly steady among non-live-in partners, though decreasing in older ages (Figure 3.3f) across female respondent age, the difference in partner age appears to greatly increase among male respondents as male respondent age increases (Figure 3.3e).

Figure 3.5 shows the mixing matrices by age for the egocentric partnership data stratified by partnership type and gender. Based on a visual assessment, the matrices are relatively similar regardless of partnership type, though there might be a bit more spread in age between older respondents with their non-live-in girlfriend/boyfriend as compared to live-in partners (Figure 3.5). Figure S3.2 shows the egocentric age mixing matrices stratified by HIV status and sex of the respondent. Using the Mantel test, all eight of these egocentric age mixing matrices (in combinations of 2) are more similar to one another than expected by random permutations.

In Figure S3.3 we display the probability of using a condom at last sex by age of participant and age of partner. While the overall matrix (Figure S3.3a) shows that young people are much more likely to use a condom than older respondents, when we stratify by partnership type, we see that live-in partners rarely use condoms with one another (Figure S3.3b), girlfriend/boyfriend partners who don't live together have a higher probability of condom use with one another (Figure S3.3c), and casual partners have a very high probability of using condoms with one another (Figure S3.3d). This is supported by Table S3.2, which shows that the proportion of partnerships that are live-in partners increases dramatically across both sexes (from 2.1% in 15-19 year old males to 89.2% in 60+ year old males and from 12.4% in 15-19 year old females to 93.4% in 60+ year old

females), while the proportion of non-household partnerships (girlfriend/boyfriends with whom you do not live and casual partnerships) decreases proportionally.

Table 3.2 shows assortativity coefficients for age mixing in the egocentric data. Assortativity by age is high across all of the egocentric age mixing data. When stratified by partner type, assortativity is highest among all partners (0.92 for men and 0.91 for women), next highest among live-in partners (0.90 for both sexes), and somewhat lower among casual partners (0.86 for men and 0.82 for women) and girlfriends/boyfriends with whom the respondent was not living (0.85 for men and 0.87 for women). Age assortativity was slightly higher among HIV-negative men (0.92) and women (0.93) than among HIV-positive men (0.83) and women (0.80). In sensitivity analysis among only most recent sexual partner (Table S3.3), assortativity coefficients were almost identical to the full sample.

3.5. Discussion

We find substantial evidence of assortative mixing by age, race, education, number of sexual partners, HIV status and ART status in this nationally representative survey of South Africans collected in 2012. Consistent with previous research, we find that, on average, respondents report a low number of sexual partners in the past year, though men consistently report more partners than women. Our results have implications for the design and implementation of interventions to combat the HIV epidemic, and can be directly used in mathematical models of HIV transmission to better capture the nuances of sexual mixing patterns and sexual degree, and thus HIV transmission, in South Africa.

Our results demonstrate assortativity by degree (or preferential attachment) in this sample of household partners. While numerous models assume assortativity by degree to some extent [32-35], to our knowledge, this has not been documented previously in South Africa. Assortative mixing by degree has been shown to have a substantial impact on STI transmission potential in simulation studies [84]. It is possible that our estimate of assortativity by degree is biased upward by an increased willingness on the part of couples in open relationships to both disclose outside partners, while individuals in closed relationships might have non-assortative outside partnerships but be unwilling to disclose these partners due to the nature of their relationship. However, we believe that this would merely dilute the effect observed, not nullify it. It is important to note that due to the very small number of household partners (1.7% of female partners and 5.4% of male partners) reporting 2+ sexual partners in this sample, our finding is based on relatively small numbers (15 couples). Additionally, we only observe this effect in household partnerships, not in non-household relationships. Future studies should attempt to capture assortativity by degree among less permanent relationships to improve our understanding of assortativity by degree in South Africa.

We find substantially higher education assortativity in the household partnership data from SABSSM IV (assortativity coefficient = 0.68) than was recently found in DHS data from 7 countries in SSA (assortativity coefficient range = 0.09-0.44) [85]. This likely is partially explained by the DHS analyses using a continuous measure of education (number of years of education) while we use a categorical measure of education (three levels of primary, secondary, and tertiary) due to limitations of the questionnaire administered. Within our results, we found that assortativity coefficients

are higher when variables are categorized rather than analyzed continuously, though South African social structure may inherently be different than other countries in SSA.

Racial assortativity was extremely high (assortativity coefficient of 0.96). Given that the South African population is largely (~80%) black African[86], this finding may not influence how we conceptualize sexual networks in South Africa. However, this level of assortativity does lead to some understanding of why South Africa has continued to maintain disparate HIV prevalence rates by race across many years of its epidemic (overall HIV prevalence in 2012 by race: black African, 15.0%; white, 0.3%; coloured, 3.1%; Indian/Asian, 0.8%) [25].

Our analyses of age mixing showed that while males are on average older than their female partners by just under 4 years, difference between partner ages varies a fair amount by male age, though seems to be more consistent by female age among non-live-in relationships. Our result for male age is consistent with a recent analysis of the South African National Communication Survey [87], and an analysis of age-mixing on Likoma Island, Malawi[88].

In the SABSSM IV data, women report a much lower proportion of sexual partners than men, a finding consistent with previous studies [58, 89-92]. It is unclear if this is due to men over-reporting number of sexual partners, women under-reporting number of sexual partners, women with higher risk (such as female sex workers) not being captured in the survey, or a combination of these biases. The discretized q-exponential distribution proposed by Handcock[83] best fit our degree data for non-household partnerships. The q-exponential distribution is an exponential family distribution with two parameters which, depending on the value of those parameters, can

characterize a long right tail [93]. Our best fit parameters for the q-exponential have a long right tail for both male and female non-household partners, implying that a strategy to reach the individuals with the greatest number of partners with combination interventions is vital to control the HIV epidemic[60]. The provision of our best fit parameters allows mathematical modelers to directly generate data from the best fit distribution to simulate data consistent with the SABSSM IV results. We observe substantial digit preference in the degree distribution data, with certain numbers (10, 20, 25, 30, 40, 50) having notably higher proportions reporting than the digits just before and after (i.e. 9 and 11). This phenomenon has been documented previously [94-96], and while it may have impacted the goodness of fit of our degree distribution to assessed probability distributions, we anticipate it does not impact the overall soundness of our results. We find a very low overall proportion of individuals who reported paying for sex in the previous year compared to other nationally representative surveys which found a prevalence closer to 10%[90], though these surveys asked about exchanging "gifts or money" and the HSRC survey was worded as a partner "paid" for sex.

The primary limitation to our analysis is the potential for reporting bias in the number of sexual partners and egocentric partner age data. We rely exclusively on self-reported sexual behavior data in an interviewer-administered survey. Previous work has shown sexual behavior reports are strongly influenced by survey methodology [92, 97-100], and that individuals often misreport their partners' ages[101, 102]. However, these studies show misreporting trends in both directions. The household partnership mixing data is less vulnerable to this potential bias since we are able to directly link participant surveys. We find high accuracy of household partner age reporting in most household

partnerships (appendix). We also are limited by grouping a range of partnership types into "non-household" partnerships (girlfriend/boyfriend, casual, paid, other), some of which (girlfriend/boyfriend) encompass a range of long-term and short-term partnering. A further limitation is that we assume that selection of secondary partners is independent of selection of primary partners, and ignore variability in the estimation of sampling weights. As such, our results likely slightly overestimate our confidence around presented assortativity coefficients. However, our sensitivity analyses find that when conducting these analyses among only most recent partner (eliminating potential dependence within a given individual's partners) our estimates remain stable and confidence intervals remain effectively the same width. We expect that even with the additional potential variability introduced by incorporating variability in sampling weights, these results would still point to substantial assortativity. An additional limitation is that we do not analyze same-sex partnerships, as a result of the potential for bias to skew our results. The inclusion of same-sex partnerships and sex workers in future research is important to fully understand patterns of sexual mixing in South Africa, but ultimately this was beyond the scope of this household survey design. Additionally, the cross-sectional nature of our findings precludes an assessment of whether HIV assortativity is indicative of an infectious process or sero-sorting.

In conclusion, our results can be utilized to improve our understanding of sexual HIV transmission in South Africa. Future mathematical modeling studies can utilize these results to appropriately simulate sexual network mixing and degree distributions. Simulation studies increasingly find that sexual network structure and mixing patterns may influence HIV prevention interventions[23, 24], thereby it is necessary to include

empirical network data in mathematical modeling studies to improve predictions of potential impact. Further elucidation of HIV transmission networks will allow for us to better target our ever-growing toolbox of HIV preventive interventions in order to achieve an AIDS-free generation.

3.6. References

1. Karim SS, Churchyard GJ, Karim QA, Lawn SD. HIV infection and tuberculosis in South Africa: an urgent need to escalate the public health response. Lancet. 2009;374(9693):921-33.

2. UNAIDS. The Gap Report. 2014. Available from: http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2014/UNAIDS_Gap_report_en.pdf. Accessed 14 Aug 2016.

3. Delva W, Leventhal GE, Helleringer S. Connecting the dots: network data and models in HIV epidemiology. AIDS. 2016;30(13):2009-20.

4. Morris M, International Union for the Scientific Study of Population. Network epidemiology : a handbook for survey design and data collection. Oxford ; New York: Oxford University Press; 2004. xii, 237 p. p.

5. Helleringer S, Mkandawire J, Kalilani-Phiri L, Kohler H-PP. Cohort Profile: The Likoma Network Study (LNS). International journal of epidemiology. 2014;43(2):545-57.

6. Helleringer S, Kohler H-PP, Chimbiri A, Chatonda P, Mkandawire J. The Likoma Network Study: Context, data collection, and initial results. Demographic research. 2009;21:427-68.

7. Helleringer S, Kohler HP. Sexual network structure and the spread of HIV in Africa: evidence from Likoma Island, Malawi. AIDS. 2007;21(17):2323-32.

8. Leclerc-Madlala S. Age-disparate and intergenerational sex in southern Africa: the dynamics of hypervulnerability. AIDS (London, England). 2008;22 Suppl 4:25.

9. Street RAA, 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. 2015;27(9):746-52.

10. Harling G, Newell M-LL, Tanser F, Kawachi I, Subramanian SV, Bärnighausen T. 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 (1999). 2014;66(4):443-51.

11. Balkus JE, Nair G, Montgomery ET, et al. Age-Disparate Partnerships and Risk of HIV-1 Acquisition Among South African Women Participating in the VOICE Trial. Journal of acquired immune deficiency syndromes (1999). 2015;70(2):212-7.

12. Gregson S, Nyamukapa CA, Garnett GP, et al. Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. Lancet (London, England). 2002;359(9321):1896-903.

13. Chapman R, White RG, Shafer LA, et al. Do behavioural differences help to explain variations in HIV prevalence in adolescents in sub‐Saharan Africa? Tropical Medicine & International Health. 2010;15(5):554-66.

14. Hallett TB, Gregson S, Lewis JJC, Lopman BA. Behaviour change in generalised HIV epidemics: impact of reducing cross-generational sex and delaying age at sexual debut. Sexually transmitted Infect. 2007;83(Suppl 1):i50-4.

15. Kelly RJ, Gray RH, Sewankambo NK, et al. Age differences in sexual partners and risk of HIV-1 infection in rural Uganda. Journal of acquired immune deficiency syndromes (1999). 2003;32(4):446-51.

16. Garnett GP, Anderson RM. Factors controlling the spread of HIV in heterosexual communities in developing countries: patterns of mixing between different age and sexual activity classes. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 1993;342(1300):137-59.

17. Wyrod R, Fritz K, Woelk G, et al. Beyond Sugar Daddies: Intergenerational Sex and AIDS in Urban Zimbabwe. AIDS and Behavior. 2010;15(6):1275-82.

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. The European journal of contraception & reproductive health care : the official journal of the European Society of Contraception. 2012;17(1):30-9.

19. Harrison A, Colvin CJ, Kuo C, Swartz A, Lurie M. Sustained High HIV Incidence in Young Women in Southern Africa: Social, Behavioral, and Structural Factors and Emerging Intervention Approaches. Current HIV/AIDS reports. 2015;12(2):207-15.

20. Maughan-Brown B, Kenyon C, Lurie MN. Partner age differences and concurrency in South Africa: implications for HIV-infection risk among young women. AIDS and Behavior. 2014;18(12):2469-76.

21. Anderson RM, Gupta S, Ng W. The significance of sexual partner contact networks for the transmission dynamics of HIV. J Acquir Immune Defic Syndr. 1990;3(4):417-29.

22. Gupta S, Anderson RM, May RM. Networks of sexual contacts: implications for the pattern of spread of HIV. AIDS. 1989;3(12):807-17.

23. Delva W, Helleringer S. Beyond Risk Compensation: Clusters of Antiretroviral Treatment (ART) Users in Sexual Networks Can Modify the Impact of ART on HIV Incidence. PLoS One. 2016;11(9):e0163159.

24. Jenness SM, Goodreau SM, Morris M, Cassels S. Effectiveness of combination packages for HIV-1 prevention in sub-Saharan Africa depends on partnership network structure: a mathematical modelling study. Sex Transm Infect. 2016;92(8):619-24.

25. Shisana O, Rehle T, Simbayi L, et al. South African National HIV Prevalence, Incidence and Behaviour Survey, 2012. Cape Town: HSRC Press; 2014.

26. Eaton JW, Bacaër N, Bershteyn A, et al. Assessment of epidemic projections using recent HIV survey data in South Africa: a validation analysis of ten mathematical models of HIV epidemiology in the antiretroviral therapy era. The Lancet Global health. 2015;3(10):608.

27. Bacaër N, Pretorius C, Auvert B. An Age-Structured Model for the Potential Impact of Generalized Access to Antiretrovirals on the South African HIV Epidemic. Bulletin of Mathematical Biology. 2010;72(8):2180-98.

28. Klein DJ, Bershteyn A, Eckhoff PA. Dropout and re-enrollment: implications for epidemiological projections of treatment programs. Aids. 2014;28.

29. Johnson L, Dorrington R, Bradshaw D. Sexual behaviour patterns in South Africa and their association with the spread of HIV: insights from a mathematical model. Demographic Research. 2009;21(11):289-340.

30. Phillips AN, Pillay D, Garnett G, et al. Effect on transmission of HIV-1 resistance of timing of implementation of viral load monitoring to determine switches from first to second-line antiretroviral regimens in resource-limited settings. Aids. 2011;25(6):843-50.

31. Cambiano V, Bertagnolio S, Jordan MR. Transmission of drug resistant HIV and its potential impact on mortality and treatment outcomes in resource-limited settings. J Infect Dis. 2013;207(Suppl 2):S57-62.

32. Dorrington R, Johnson L, Budlender D. ASSA2008 AIDS and demographic models: user guide. Cape Town: ASSA; 2010. Accessed 15 Aug 2015.

33. Eaton JW, Hallett TB. Why the proportion of transmission during early-stage HIV infection does not predict the long-term impact of treatment on HIV incidence. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(45):16202-7.

34. Cori A, Ayles H, Beyers N, et al. HPTN 071 (PopART): a cluster-randomized trial of the population impact of an HIV combination prevention intervention including universal testing and treatment: mathematical model. PloS one. 2014;9(1).

35. Johnson LF, Hallett TB, Rehle TM, Dorrington RE. The effect of changes in condom usage and antiretroviral treatment coverage on human immunodeficiency virus incidence in South Africa: a model-based analysis. Journal of the Royal Society, Interface / the Royal Society. 2012;9(72):1544-54.

36. Hontelez JA, Lurie MN, Bärnighausen T, et al. Elimination of HIV in South Africa through expanded access to antiretroviral therapy: a model comparison study. PLoS medicine. 2013;10(10).

37. Eaton JW, Johnson LF, Salomon JA, et al. HIV treatment as prevention: systematic comparison of mathematical models of the potential impact of antiretroviral therapy on HIV incidence in South Africa. PLoS Med. 2012;9(7):e1001245.

38. Abuelezam NN, McCormick AW, Fussell T, et al. Can the Heterosexual HIV Epidemic be Eliminated in South Africa Using Combination Prevention? A Modeling Analysis. Am J Epidemiol. 2016;184(3):239-48.

39. Department of Health Republic of South Africa. South Africa Demographic and Health Survey 1998: Full Report. 1998. Available from: http://dhsprogram.com/pubs/pdf/FR131/FR131.pdf. Accessed 12 March 2016.

40. Ott MQ, Bärnighausen T, Tanser F, Lurie MN, Newell M-LL. Age-gaps in sexual partnerships: seeing beyond 'sugar daddies'. AIDS (London, England). 2011;25(6):861-3.

41. Lurie MN, Williams BG, Zuma K, 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. Bohl DD, McFarland W, Raymond HF. Improved measures of racial mixing among men who have sex with men using Newman's assortativity coefficient. Sexually transmitted infections. 2011;87(7):616-20.

43. Fujimoto K, Williams ML. Racial/Ethnic Differences in Sexual Network Mixing: A Log-Linear Analysis of HIV Status by Partnership and Sexual Behavior Among Most at-Risk MSM. AIDS and behavior. 2015;19(6):996-1004.

44. Doherty IA, Schoenbach VJ, Adimora AA. Sexual Mixing Patterns and Heterosexual HIV Transmission Among African Americans in the Southeastern United States. JAIDS Journal of Acquired Immune Deficiency Syndromes. 2009;52(1):114.

45. Schneider JA, Cornwell B, Ostrow D, et al. Network mixing and network influences most linked to HIV infection and risk behavior in the HIV epidemic among black men who have sex with men. American journal of public health. 2013;103(1):36.

46. Chow EP, Read TR, Law MG, Chen MY, Bradshaw CS, Fairley CK. Assortative sexual mixing patterns in male?female and male?male partnerships in Melbourne, Australia: implications for HIV and sexually transmissible infection transmission. Sex Health. 2016.

47. Chow EP, Fairley CK. Assortative sexual mixing among heterosexuals in Australia: implications for herd protection in males from a female human papillomavirus vaccination program. Sex Health. 2016.

48. Doherty IA, Adimora AA, Muth SQ, Serre ML, Leone PA, Miller WC. Comparison of sexual mixing patterns for syphilis in endemic and outbreak settings. Sex Transm Dis. 2011;38(5):378-84.

49. Rothenberg R, Muth SQ. Large-network concepts and small-network characteristics: fixed and variable factors. Sex Transm Dis. 2007;34(8):604-12.

50. Wilson EC, Santos GM, Raymond HF. Sexual mixing and the risk environment of sexually active transgender women: data from a respondent-driven sampling study of HIV risk among transwomen in San Francisco, 2010. BMC Infect Dis. 2014;14:430.

51. Sudhinaraset M, Raymond HF, McFarland W. Convergence of HIV prevalence and inter-racial sexual mixing among men who have sex with men, San Francisco, 2004-2011. AIDS Behav. 2013;17(4):1550-6.

52. Prah P, Copas AJ, Mercer CH, Nardone A, Johnson AM. Patterns of sexual mixing with respect to social, health and sexual characteristics among heterosexual couples in England: analyses of probability sample survey data. Epidemiol Infect. 2015;143(7):1500-10.

53. Tieu HV, Liu TY, Hussen S, et al. Sexual Networks and HIV Risk among Black Men Who Have Sex with Men in 6 U.S. Cities. PLoS One. 2015;10(8):e0134085.

54. Mahapatra B, Lowndes CM, Gurav K, et al. Degree and correlates of sexual mixing in female sex workers in Karnataka, India. Sex Health. 2013;10(4):305-10.

55. Birkett M, Kuhns LM, Latkin C, Muth S, Mustanski B. The sexual networks of racially diverse young men who have sex with men. Arch Sex Behav. 2015;44(7):1787-97.

56. Harling G, Barnighausen T. The role of partners' educational attainment in the association between HIV and education amongst women in seven sub-Saharan African countries. J Int AIDS Soc. 2016;19(1):20038.

57. Kenyon C, Colebunders R. Birds of a feather: homophily and sexual network structure in sub-Saharan Africa. International journal of STD & AIDS. 2013;24(3):211-5.

58. 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.

59. Morris M. Telling tails explain the discrepancy in sexual partner reports. Nature. 1993;365(6445):437-40.

60. Anderson RM, May RM. Infectious diseases of humans : dynamics and control. Oxford ; New York: Oxford University Press; 1991. viii, 757 p. p.

61. Newman ME. Spread of epidemic disease on networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2002;66(1 Pt 2):016128.

62. Boguna M, Pastor-Satorras R, Vespignani A. Absence of epidemic threshold in scale-free networks with degree correlations. Phys Rev Lett. 2003;90(2):028701.

63. Liljeros F, Edling CR, Amaral LA, Stanley EH, Åberg Y. The web of human sexual contacts. Nature. 2001;411(6840):907-8.

64. Schneeberger A, Mercer CH, Gregson SA, et al. Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. Sex Transm Dis. 2004;31(6):380-7.

65. Latora V, Nyamba A, Simpore J, et al. Network of sexual contacts and sexually transmitted HIV infection in Burkina Faso. J Med Virol. 2006;78(6):724-9.

66. Jones JH, Handcock. An assessment of preferential attachment as a mechanism for human sexual network formation. Proceedings of the Royal Society B: Biological Sciences. 2003;270(1520):11231128.

67. Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. arXiv. 2009;0706.1062.

68. Krivitsky PN, Handcock MS, Raftery AE, Hoff PD. Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models. Social networks. 2009.

69. Hamilton DT, Handcock MS, Morris M. Degree distributions in sexual networks: a framework for evaluating evidence. Sexually transmitted diseases. 2008;35(1):30-40.

70. Handcock MS, Jones JH. Likelihood-based inference for stochastic models of sexual network formation. Theor Popul Biol. 2004;65(4):413-22.

71. Handcock MS, Jones JH. Interval estimates for epidemic thresholds in two-sex network models. Theor Popul Biol. 2006;70(2):125-34.

72. Garnett GP. Inferring mechanisms for sexual partnership formation from the distribution of sexual partner numbers. Sex Transm Dis. 2008;35(1):41-2.

73. Dezso Z, Barabasi AL. Halting viruses in scale-free networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2002;65(5 Pt 2):055103.

74. Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. Proceedings Biological sciences / The Royal Society. 2014;281(1785):20140268.

75. Maslov S, Sneppen K. Specificity and stability in topology of protein networks. Science (New York, NY). 2002;296(5569):910-3.

76. Mantel N. The detection of disease clustering and a generalized regression approach. Cancer research. 1967.

77. Newman MEJ. Mixing patterns in networks. Physical Review E. 2003.

78. Newman MEJ. Assortative mixing in networks. Assortative mixing in networks. 2002.

79. Fisher RA. Frequency distribution o the values of the correlation coefficient in samples of an indefinitely large population. Biometrika. 1915;10(4):507-21.

80. Young SK, Lyles RH, Kupper LL, Keys JR, Martin SL, Costenbader EC. Assortativity coefficient-based estimation of population patterns of sexual mixing when cluster size is informative. Sex Transm Infect. 2014;90(4):332-6.

81. Latora V, Nyamba A, Simpore J, et al. Network of sexual contacts and sexually transmitted HIV infection in Burkina Faso. Journal of medical virology. 2006;78(6):724-9.

82. Handcock MS, Jones J. Likelihood-based inference for stochastic models of sexual network formation. Theoretical Population Biology. 2004;65(4):413422.

83. Handcock MS. degreenet: Models for Skewed Count Distributions Relevant to Networks. Los Angeles, CA2015.

84. Doherty IA, Shiboski S, Ellen JM, Adimora AA, Padian NS. Sexual bridging socially and over time: a simulation model exploring the relative effects of mixing and concurrency on viral sexually transmitted infection transmission. Sex Transm Dis. 2006;33(6):368-73.

85. Harling G, Bärnighausen T. The role of partners' educational attainment in the association between HIV and education amongst women in seven sub-Saharan African countries. Journal of the International AIDS Society. 2016;19(1):20038.

86. Statistics South Africa. Census 2011. South Africa 2012. Available from: http://www.statssa.gov.za/Publications/P03014/P030142011.pdf. Accessed 2 Aug 2014.

87. 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(12):2469-76.

88. 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;6:36121.

89. Wellings K, Collumbien M, Slaymaker E, et al. Sexual behaviour in context: a global perspective. Lancet. 2006;368(9548):1706-28.

90. Carael M, Cleland J, Deheneffe JC, Ferry B, Ingham R. Sexual behaviour in developing countries: implications for HIV control. AIDS. 1995;9(10):1171-5.

91. Curtis SL, Sutherland EG. Measuring sexual behaviour in the era of HIV/AIDS: the experience of Demographic and Health Surveys and similar enquiries. Sex Transm Infect. 2004;80 Suppl 2:ii22-7.

92. 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.

93. Jr. SP, Mendes RS, Malacarne LC. q-exponential, Weibull, and q-Weibull distributions: an empirical analysis. Physica A. 2003;324:678-88.

94. Foxman B, Newman M, Percha B, Holmes KK, Aral SO. Measures of sexual partnerships: lengths, gaps, overlaps, and sexually transmitted infection. Sexually transmitted diseases. 2006;33(4):209-14.

95. Roberts JM, Brewer DD. Measures and tests of heaping in discrete quantitative distributions. Journal of Applied Statistics. 2001;28(7):887-96.

96. Mikolajczyk RT, Kretzschmar M. Collecting social contact data in the context of disease transmission: prospective and retrospective study designs. Social Networks. 2008.

97. 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:ii43-8.

98. Cleland J, Boerma JT, Carael M, Weir SS. Monitoring sexual behaviour in general populations: a synthesis of lessons of the past decade. Sex Transm Infect. 2004;80 Suppl 2:ii1-7.

99. Helleringer S, Kohler HP, Kalilani-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.

100. Mensch BS, Hewett PC, Erulkar AS. The reporting of sensitive behavior by adolescents: a methodological experiment in Kenya. Demography. 2003;40(2):247-68.

101. Harling G, Tanser F, Mutevedzi T, Barnighausen T. Assessing the validity of respondents' reports of their partners' ages in a rural South African population-based cohort. BMJ Open. 2015;5(3):e005638.

102. 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.

Table 3.1 Description of household data, percent in-group and relative increase in like-with-like mixing relative to random mixing

Characteristic	% (n/N) or mean (sd) of total household partner sample	% in-group*	Relative increase compared to random mixing (95% bootstrapped CIs)
Males			
Race			
Black	63.1 (2019/5110)	98.7 (1962/2018)	1.56 (1.53-1.60)
Coloured	13.1 (1131/5110)	93.1 (1087/1130)	7.20 (6.74-7.72)
Indian/Asian	4.0 (942/5110)	95.7 (917/942)	24.08 (21.25-27.68)
White	19.6 (1006/5110)	98.4 (987/1005)	4.96 (4.70-5.27)
Education			
Primary or less	31.2 (1122/4719)	71.2 (746/1085)	2.58 (2.47-2.69)
Some secondary or secondary completion	54.3 (2892/4719)	82.9 (2398/2842)	1.42 (1.39-1.45)
Some tertiary or higher	14.5 (705/4719)	67.6 (441/691)	4.84 (4.53-5.17)
Employment status			
Unemployed	41.7 (1488/4258)	79.3 (1130/1416)	1.29 (1.26-1.32)
Employed	58.3 (2770/4258)	50.6 (1264/2565)	1.32 (1.28-1.36)
HIV Status			
Positive	10.3 (252/3148)	68.1 (162/234)	4.96 (4.53-5.38)
Negative	89.7 (2896/3148)	92.3 (2467/2612)	1.07 (1.06-1.08)
Number of lifetime sexual partners			
1	20.0 (1036/3903)	80.6 (797/932)	1.94 (1.85-2.03)
2-3	29.5 (1234/3903)	43.9 (466/1130)	1.06 (1.00-1.12)
4-5	21.2 (761/3903)	17.4 (100/693)	1.56 (1.34-1.77)
6+	29.4 (872/3903)	12.3 (90/791)	2.08 (1.83-2.35)
Number of sexual partners in past 12 mos			
0	12.8 (543/4191)	83.6 (409/505)	6.62 (6.18-7.12)
1	81.8 (3499/4191)	95.5 (3146/3257)	1.11 (1.10-1.13)
2+	5.4 (149/4191)	10.6 (15/139)	6.41 (4.20-8.63)
HIV Testing History			
Never	34.4 (1419/4332)	47.5 (659/1346)	1.95 (1.86-2.04)
2+ years ago	13.8 (669/4332)	36.8 (221/628)	2.09 (1.90-2.30)
1-2 years ago	11.3 (500/4332)	24.7 (119/475)	2.09 (1.79-2.40)
Within past year	40.5 (1744/4332)	64.2 (1011/1630)	1.39 (1.35-1.44)
ART Status (among those living with HIV)			
On ART	29.9 (66/250)	71.0 (34/54)	2.27 (1.92-2.62)
Not on ART	70.1 (184/250)	88.1 (91/108)	1.28 (1.19-1.37)
Females			
Race			
Black	63.1 (2021/5122)	98.5 (1962/2017)	1.56 (1.53-1.60)
Coloured	13.0 (1156/5122)	94.2 (1087/1151)	7.17 (6.72-7.67)
Indian/Asian	4.0 (932/5122)	97.6 (917/928)	24.07 (21.46-27.44)
White	19.8 (1008/5122)	97.1 (987/1006)	4.96 (4.72-5.23)

Education			
Primary or less	27.6 (1076/4772)	80.0 (746/1028)	2.56 (2.47-2.66)
Some secondary or secondary completion	58.4 (3058/4772)	77.3 (2398/2968)	1.42 (1.39-1.45)
Some tertiary or higher	14.1 (638/4772)	70.3 (441/622)	4.85 (4.55-5.19)
Employment status			
Unemployed	61.6 (1863/4695)	55.1 (1130/2431)	1.32 (1.29-1.35)
Employed	38.4 (2832/4695)	76.4 (1264/1550)	1.31 (1.28-1.35)
HIV Status			
Positive	13.7 (388/3626)	51.4 (162/307)	5.00 (4.57-5.49)
Negative	86.3 (3238/3626)	96.0 (2467/2539)	1.07 (1.06-1.08)
Number of lifetime sexual partners			
1	41.7 (2226/4459)	37.9 (797/1784)	1.89 (1.80-1.99)
2-3	41.3 (1593/4459)	32.1 (466/1254)	1.09 (1.03-1.17)
4-5	11.1 (428/4459)	33.2 (100/339)	1.58 (1.39-1.80)
6+	5.9 (212/4459)	57.1 (90/169)	1.93 (1.71-2.16)
Number of sexual partners in past 12 mos			
0	12.7 (577/4624)	80.0 (409/501)	6.28 (5.86-6.80)
1	85.6 (4002/4624)	92.1 (3146/3359)	1.13 (1.11-1.14)
2+	1.7 (45/4624)	31.6 (15/41)	5.87 (3.89-7.95)
HIV Testing History			
Never	24.4 (1237/4753)	63.9 (659/1108)	1.87 (1.78-1.96)
2+ years ago	17.3 (878/4753)	29.7 (221/736)	2.14 (1.94-2.34)
1-2 years ago	11.8 (566/4753)	24.2 (119/480)	2.14 (1.83-2.46)
In past year	46.4 (2072/4753)	56.9 (1011/1755)	1.40 (1.36-1.44)
ART Status (among those living with HIV)			
On ART	31.3 (115/389)	78.4 (34/51)	2.62 (2.26-3.10)
Not on ART	68.7 (274/389)	82.3 (91/111)	1.18 (1.10-1.28)

*% In group represents the proportion of all individuals paired to a household partner who had a partner that matched their group

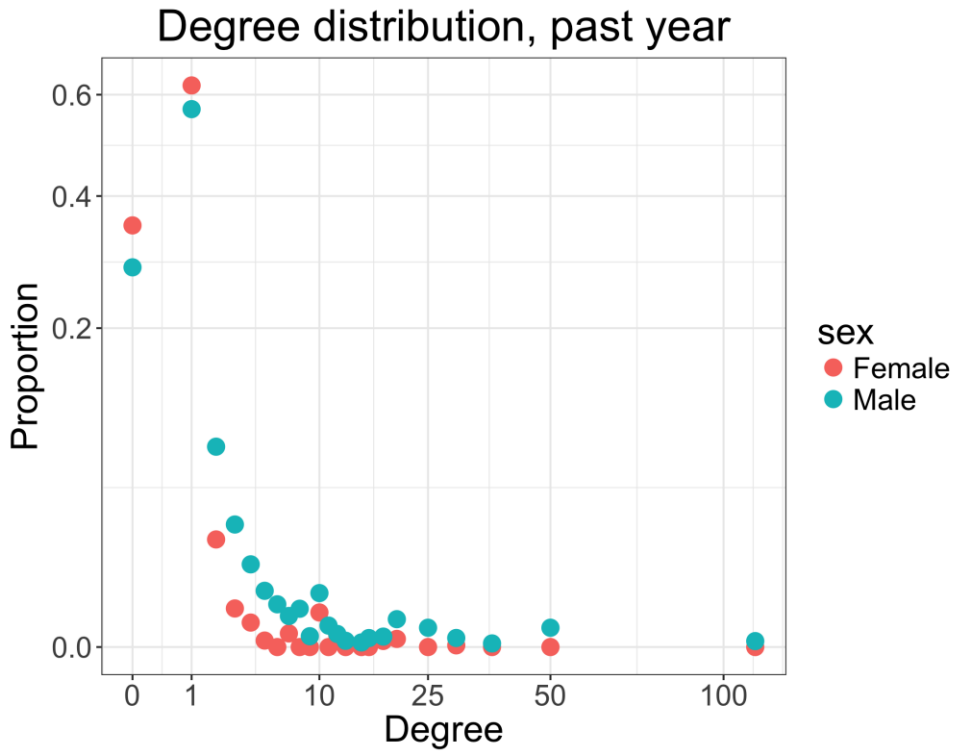
Denominators different for columns due to missing data for partners

Table 3.2: Newman's assortativity coefficients

Characteristic	Assortativity (95% pseudo-pop closed-form CI)	Assortativity (95% pseudo-pop jack-knifed CI)	Assortativity (95% weighted bootstrap CI)
Household Partnerships			
Age	0.88 (0.88-0.89)	0.88 (0.88-0.88)	0.88 (0.87-0.90)
Race	0.96 (0.95-0.98)	0.96 (0.95-0.97)	0.96 (0.95-0.97)
Education	0.68 (0.67-0.70)	0.68 (0.66-0.71)	0.68 (0.67-0.70)
Employment status	0.48 (0.46-0.50)	0.48 (0.45-0.51)	0.48 (0.46-0.50)
HIV status	0.67 (0.65-0.69)	0.67 (0.63-0.71)	0.67 (0.64-0.69)
Lifetime number of partners, categorical	0.28 (0.27-0.30)	0.28 (0.26-0.31)	0.28 (0.27-0.30)
Lifetime number of partners, continuous	0.32 (0.28-0.35)	0.32 (0.26-0.37)	0.32 (0.19-0.45)
Past year number of partners, categorical	0.72 (0.70-0.73)	0.72 (0.69-0.75)	0.72 (0.70-0.74)
Past year number of partners, continuous	0.19 (0.16-0.22)	0.19 (0.17-0.20)	0.19 (0.12-0.27)
HIV Testing History	0.43 (0.42-0.45)	0.43 (0.41-0.46)	0.43 (0.42-0.45)
ARV Status	0.70 (0.61-0.79)	0.70 (0.59-0.80)	0.70 (0.61-0.77)
Egocentric Partnerships - Men			
Age	0.919 (0.916-0.922)	0.919 (0.912-0.926)	0.919 (0.913-0.924)
Age among live-in partners	0.902 (0.895-0.907)	0.902 (0.889-0.914)	0.902 (0.892-0.910)
Age among non-live-in girlfriend/boyfriends	0.851 (0.843-0.860)	0.851 (0.830-0.874)	0.851 (0.836-0.867)
Age among casual partners	0.860 (0.838-0.879)	0.860 (0.812-0.908)	0.860 (0.800-0.904)
Age among HIV-negative	0.915 (0.911-0.919)	0.915 (0.906-0.924)	0.915 (0.914-0.928)
Age among HIV-positive	0.830 (0.809-0.848)	0.830 (0.793-0.866)	0.830 (0.795-0.860)
Egocentric Partnerships - Women			
Age	0.909 (0.905-0.913)	0.909 (0.899-0.920)	0.909 (0.902-0.916)
Age among live-in partners	0.899 (0.893-0.905)	0.899 (0.885-0.913)	0.899 (0.890-0.907)
Age among non-live-in girlfriend/boyfriends	0.868 (0.859-0.877)	0.868 (0.842-0.895)	0.868 (0.846-0.889)
Age among casual partners	0.817 (0.763-0.859)	0.817 (0.714-0.919)	0.817 (0.732-0.880)
Age among HIV-negative	0.930 (0.925-0.934)	0.930 (0.921-0.938)	0.930 (0.924-0.936)
Age among HIV-positive	0.795 (0.773-0.814)	0.795 (0.725-0.837)	0.795 (0.760-0.816)

Figure 3.1 Degree Distribution by sex for participants aged 15-65 a) in the past year and b) over respondents' lifetime, South Africa, 2012

A



B

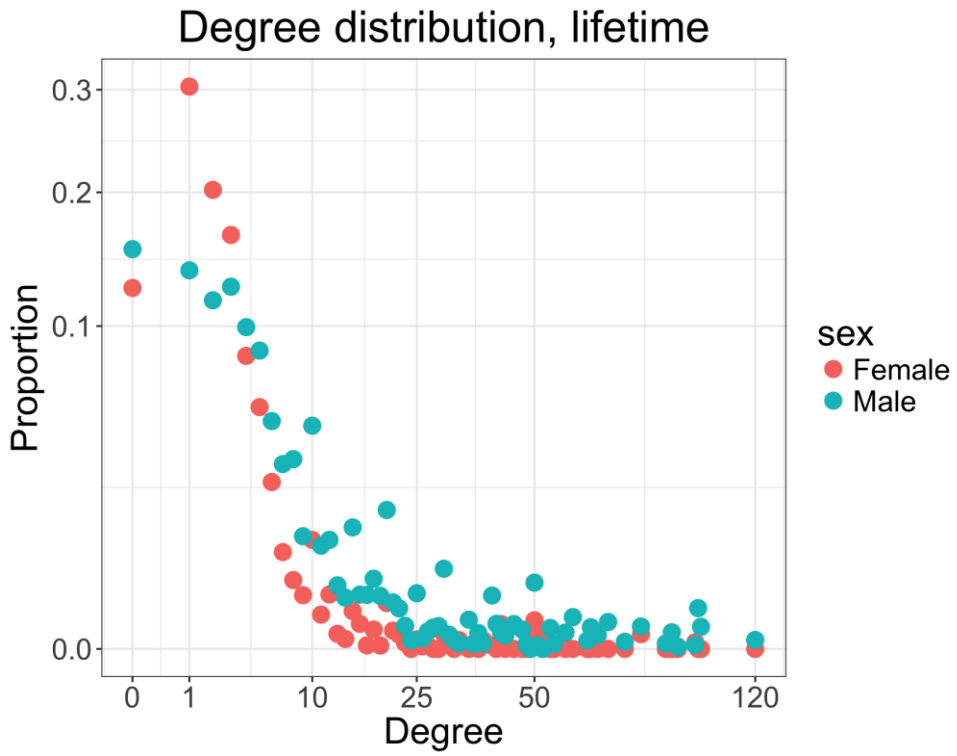


Figure 3.2 Degree Distribution by age and sex a) in the past year and b) over the respondents' lifetime

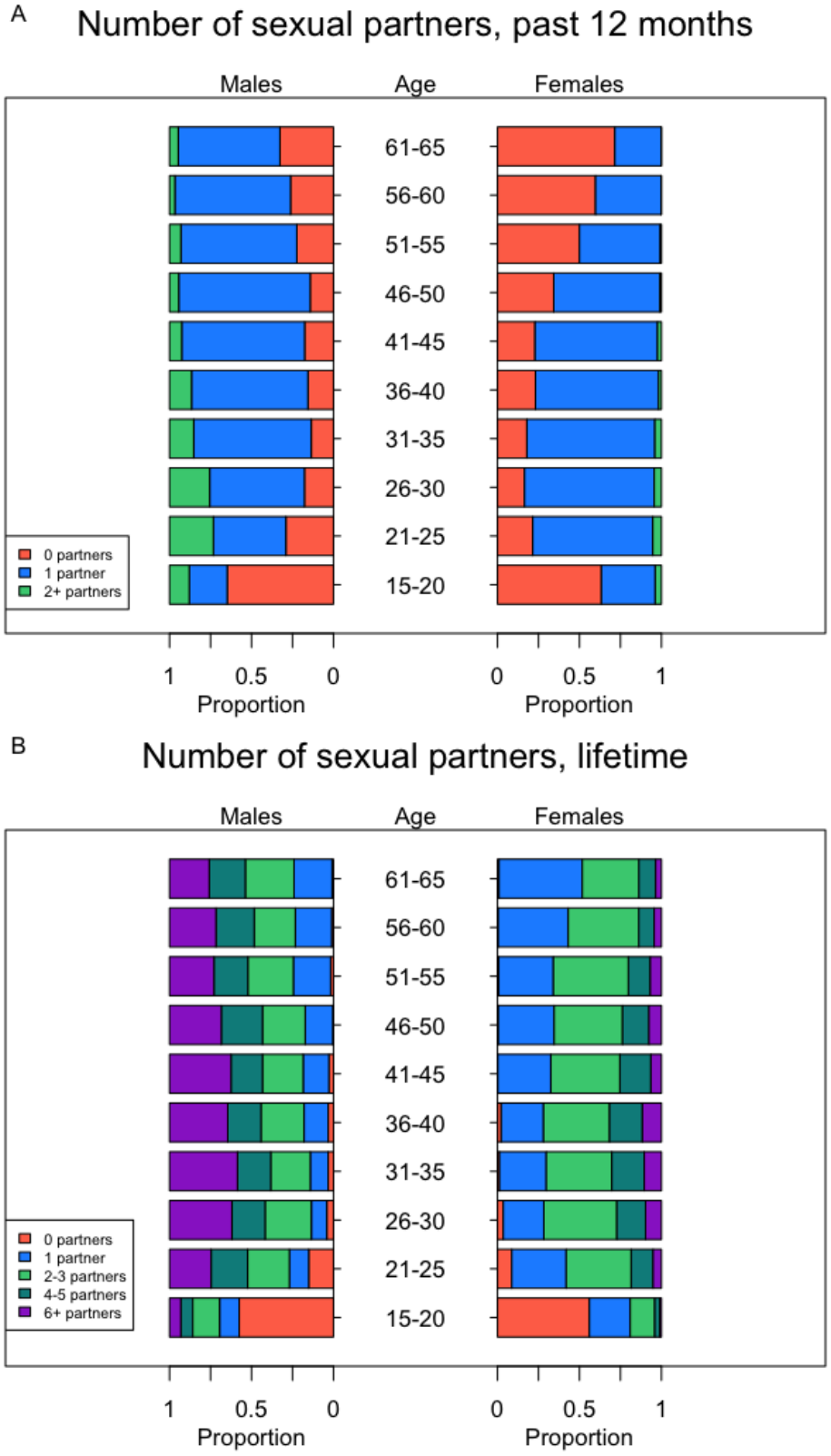


Figure 3.3: Age difference between male and female partners among household data on partners of a) males and b) females, among egocentric data on live-in partners of c) males and d) females, among egocentric data on boyfriend/girlfriend or casual partners of e) males and f) females

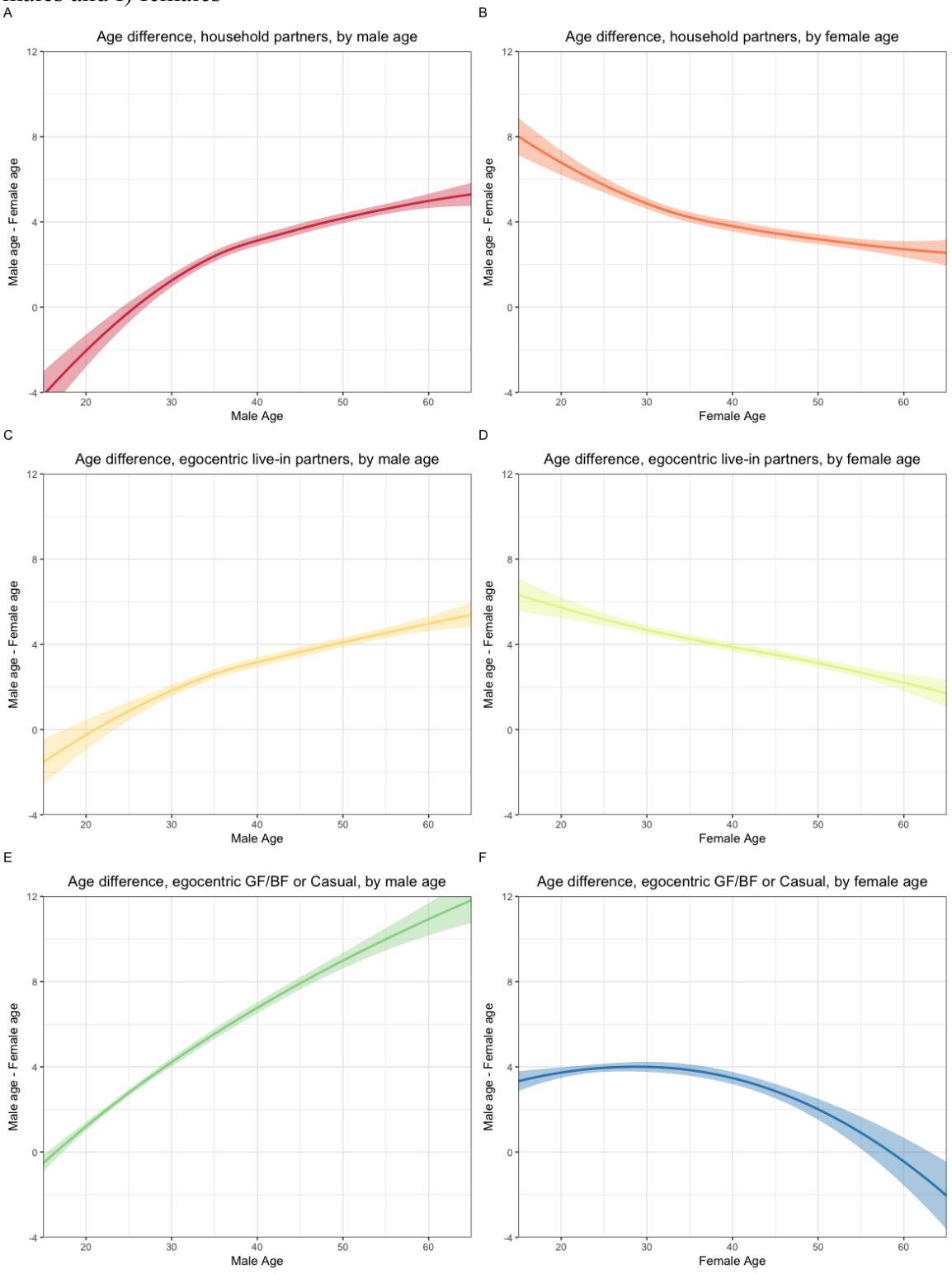
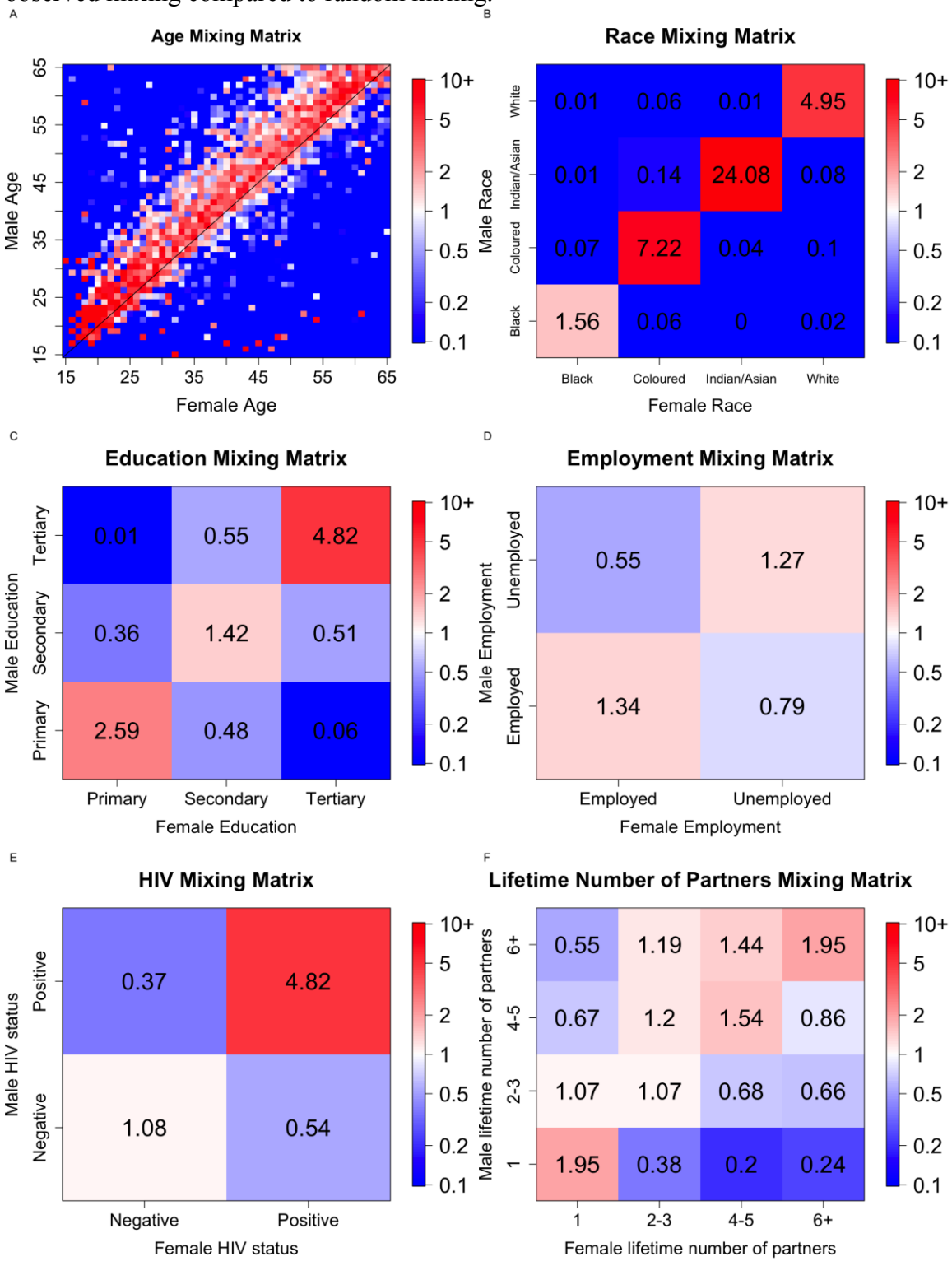
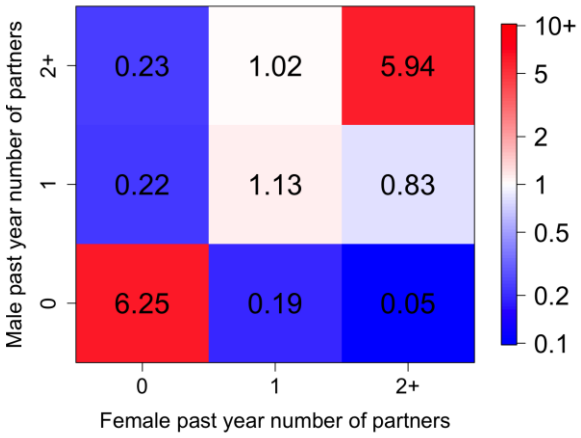


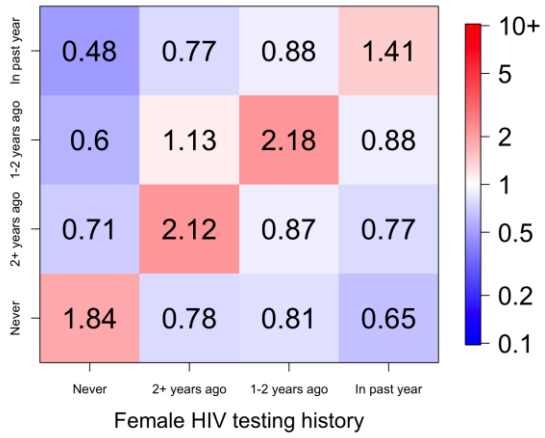
Figure 3.4 Mixing matrices for household partnership data for a) age, b) race, c) education, d) employment status, e) HIV status, f) lifetime number of sexual partners, g) number of sexual partners in past year, h) HIV testing history and i) ART status among couples where both are living with HIV. Each cell represents the relative ratio of observed mixing compared to random mixing.



G


Past Year Number of Partners Mixing Matrix
Male past year number of partners
2+
1
0
0.23
1.02
5.94
0.22
1.13
0.83
6.25
0.19
0.05
0
1
2+
Female past year number of partners
10+
5
2
1
0.5
0.2
0.1

H


HIV Testing History Mixing Matrix
Male HIV testing history
In past year
1-2 years ago
2+ years ago
Never
0.48
0.77
0.88
1.41
0.6
1.13
2.18
0.88
0.71
2.12
0.87
0.77
1.84
0.78
0.81
0.65
Never
2+ years ago
1-2 years ago
In past year
Female HIV testing history
10+
5
2
1
0.5
0.2
0.1

I

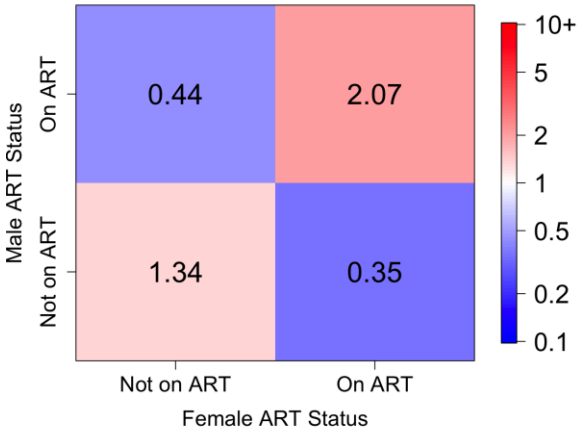
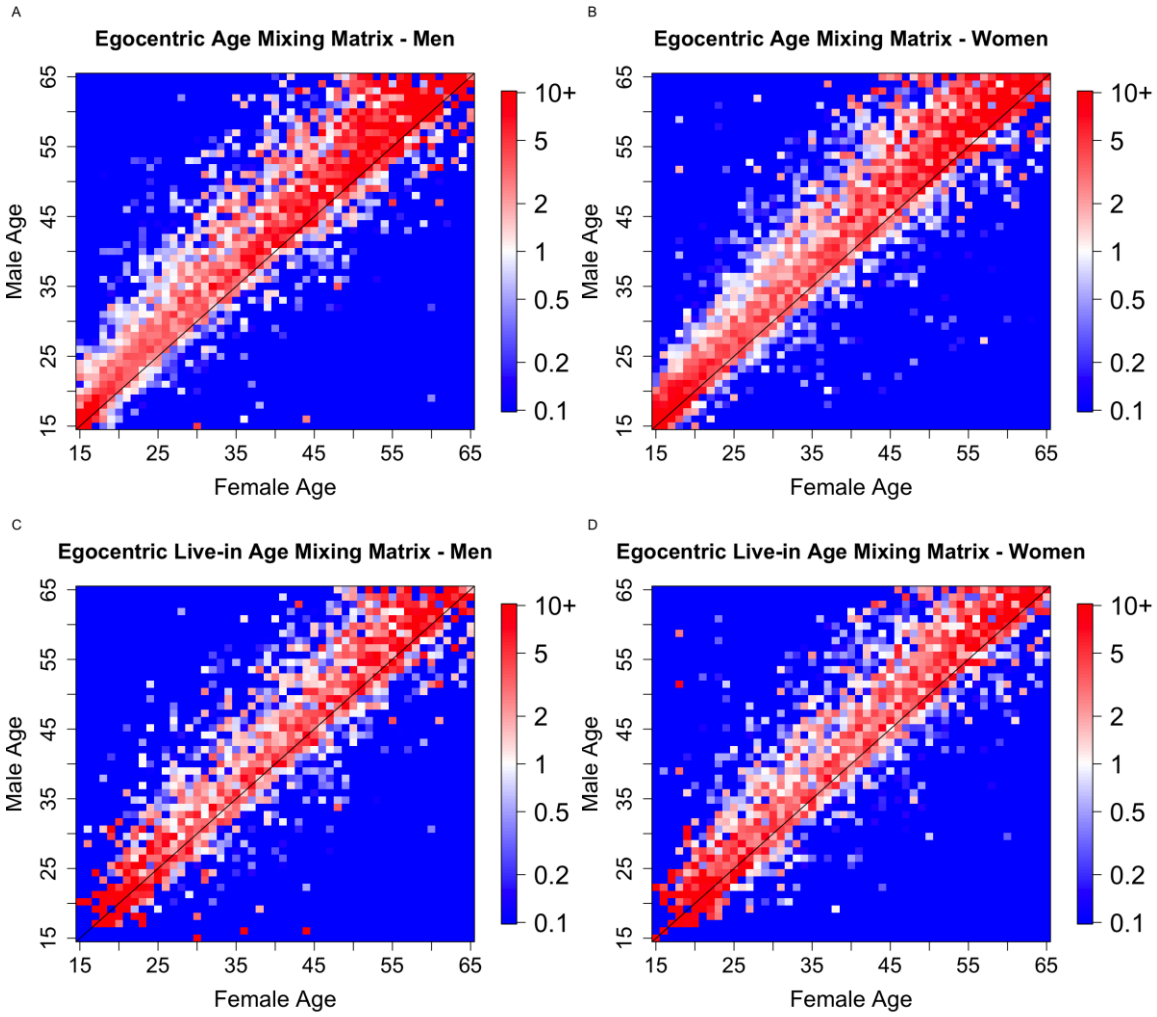
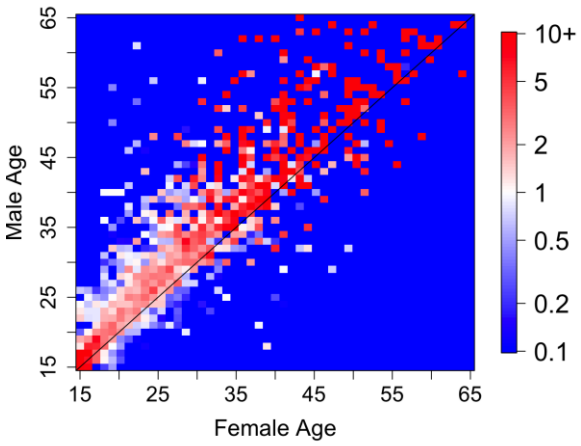

ART Mixing Matrix, among PLHIV
Male ART Status
On ART
Not on ART
0.44
2.07
1.34
0.35
Not on ART
On ART
Female ART Status
10+
5
2
1
0.5
0.2
0.1

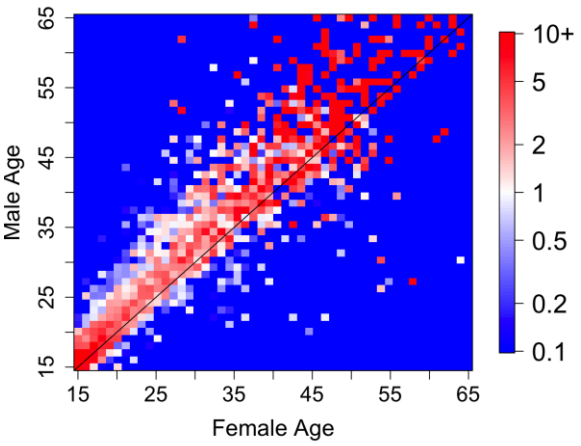
Figure 3.5: Age mixing matrix for all a) men's egocentric data, b) women's egocentric data, c) men's egocentric live-in partners, d) women's egocentric live-in partners, e) men's egocentric boyfriend/girlfriend, f) women's egocentric boyfriend/girlfriend, g) men's egocentric casual partnerships and h) women's egocentric casual partnerships. Each cell represents the relative ratio of observed mixing compared to random mixing.



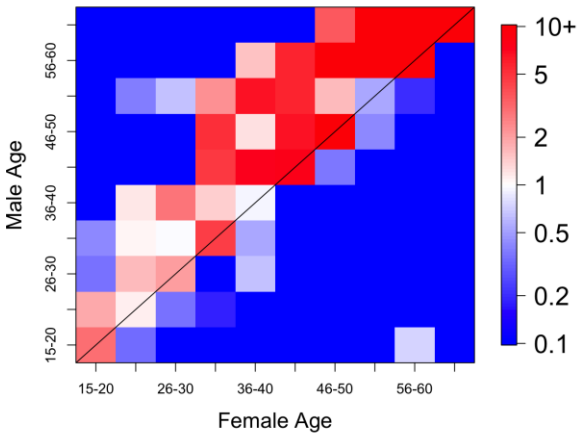
E


Egocentric GF/BF Age Mixing Matrix - Men
Male Age
Female Age
15
25
35
45
55
65
10+
5
2
1
0.5
0.2
0.1

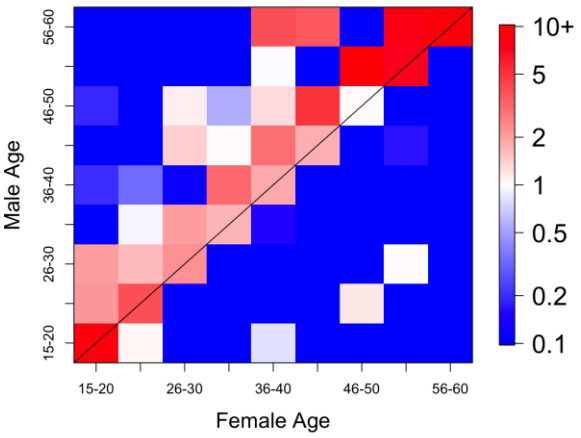
F


Egocentric GF/BF Age Mixing Matrix - Women
Male Age
Female Age
15
25
35
45
55
65
10+
5
2
1
0.5
0.2
0.1

G


Egocentric Casual Age Mixing Matrix - Men
Male Age
Female Age
15-20
26-30
36-40
46-50
56-60
10+
5
2
1
0.5
0.2
0.1

H


Egocentric Casual Age Mixing Matrix - Women
Male Age
Female Age
15-20
26-30
36-40
46-50
56-60
10+
5
2
1
0.5
0.2
0.1

3.7. Appendix

3.7.1. Weighted Pseudopopulation

To incorporate complex survey weights in mixing matrices, each individual was up- or down-weighted within mixing matrices to represent their relative contribution to the overall South African population. All weights summed to the total sample size. We refer to this mixing matrix as a weighted pseudopopulation that is (under our assumptions listed above) representative of the "true" mixing matrix across South Africa.

3.7.2. Mantel Test

The Mantel test [1] is a permutation test of the similarity between matrices wherein the rows and columns of the matrices are permuted randomly and a Z statistic is calculated, with the Z statistic between matrices X and Y being:

$$Z = \sum_{ij} X_{ij} * Y_{ij}$$

We then compare our observed value to the permuted null distribution to assess if these matrices are more similar than expected by chance.

3.7.3. Missingness Analysis

Overall, there were 24,367 participants aged 15-65 in SABSSM IV. Of these individuals, 953 (3.6%) did not respond to the question on having ever had sex, 2,431 (9.1%) did not respond to the question on number of lifetime partners, 1,252 (4.7%) did not respond to the question on sexual activity in the past year, and 1,283 (4.8%) did not respond to the question on number of sexual partners in the past year.

Of all participants, 14,422 reported the sex and age of one partner on the egocentric partner census, 567 reported sex and age for two partners, and 275 reported for three partners. We found that when we compared our egocentric partnerships with full age and sex to the number of partners reported earlier in the questionnaire (degree in the past year), 20,745 respondents had no disparity in number of partners, but 1,430 (5.6%) had a disparity between their reports, with 999 reporting one more participant in the degree question than the egocentric, 305 reporting 2 more in the degree question, and 24 reporting 3+ more in the degree question, while 88 reported one more in the egocentric section than the degree question, 11 two more in the egocentric section, and 3 three more in the egocentric section, representing 1,681 (6.6%) egocentric partnerships missing. Table S3.4 shows the degree reported among the 1,430 with a disparity between their degree distribution and egocentric data by the number difference in partners. We see that in the largest group (those with 1 more partner reported in degree than egocentric), just over half (56%) reported 1 partner the degree question but none in the egocentric section, while another third (35%) reported 2 partners in the degree distribution but only one in the egocentric section. This is a relatively low proportion overall, but given that only 1,495 participants reported two or more sexual partners in the past year on the degree question, less than half provided consistent reporting (725, 48.5%) between the degree question and the egocentric data.

We hypothesized that the partners that were accurately reported were preferentially household partnerships, and to test this compared the distribution of partnership types by number of partners, finding little difference. Among individuals who reported more than one egocentric partner, 16% of their first partners were live-in

partners, while 5% of the second partners reported were live-in and 6% of third partners were live-in partners. Among the 737 individuals who reported more partners in the degree distribution data than the egocentric data for whom there is egocentric data, 20% of their first partners were live-in, 6% of their second partners were live-in, and 7% of their third partners were live-in partners.

Since our primary analyses of the egocentric partnership data surround age of egocentric partner, we attempted to elucidate if the age of the partners reported second or later on the egocentric partnership questionnaire were systematically older or younger partners, to determine the need for IPW to account for missing partner age on second+ reported partnerships. We find overall no large difference in partners' age between first and second (median 0 years difference, IQR= -2 to 3, range= -22 to 45), second and third (median 0 years difference, IQR= -2 to 2, range = -29 to 22), and first and third (median 1 years difference, IQR= -2 to 3, range = -38 to 31).

3.7.4. Accuracy of self-reported live-in partner age

Among household partners, we assessed the accuracy of self-reported live-in partner age as compared to linked household partner age. Among male reports of their female live-in partner ages, the majority (86%) had no difference between their self-reported female age and linked female age. In 95% of live-in partner age reports, men reported a partner's age between 1 year younger and 5 years older than their female partners' age, and were within 10 years above or below in 98.4% of cases. The additional discrepancy in reports (1.6% of partner age reports) ranged from 37 years younger than partner to 43 years older than the household partner – which we cannot reliably assess between error in reporting by the respondent, error in recording by

interviewer or by data entry, or a different household partner being described in the egocentric data.

Among female reports of their male live-in partner ages, the majority (86%) had no difference between their self-reported male partner ages and linked partner age. In 95% of female partner age reports, partner age reported by females ranged from 4 years younger than household partner age to 1 year older than household partner age. In 98.7% of cases, females reported partner age within 10 years of the reported partners' age. In the 1.3% of additional cases, the difference ranged from 47 years younger than linked partner age to 43 years older.

Table S3.1: Maximum likelihood estimates for degree distributions of male and female non-household and household partnerships, including the log-likelihood, the corrected AIC, and the BIC for the model.
See appendix below for PMFs with each variable (V1, V2, V3, V4) designated

Distribution	**V1**	**V2**	**V3**	**V4**	**Log-likelihood**	**AICc**	**BIC**
Non-household Partnerships							
Females							
Discrete q-exponential	1.79	0.33	-	-	-1420	2846	2865
Waring	2.93	0.84	-	-	-1422	2849	2868
Geometric-Waring	2.89	0.84	141.85	-	-1422	2851	2877
Negative binomial-Waring	2.93	0.84	164.07	0.2	-1422	2853	2885
Negative binomial-Yule	2.75	2.41	0.03	-	-1427	2863	2888
Discrete Pareto	3.75	-	-	-	-1438	2880	2892
Negative binomial	0.24	0.25	-	-	-1444	2895	2914
Yule	8.28	-	-	-	-1527	3058	3071
Geometric	1.18	-	-	-	-1848	3700	3713
Poisson	0.18	-	-	-	-2341	4686	4698
Geometric-Yule	7.95	1104.96	-	-	-3078	6164	6190
Poisson lognormal	0.27	0.59	-	-	-3645	7295	7314
Males							
Discrete q-exponential	1.99	1.42	-	-	-5389	10783	10804
Yule	2.94	-	-	-	-5390	10783	10798
Waring	3.04	0.49	-	-	-5389	10784	10806
Negative binomial-Yule	2.94	84.82	0.1	-	-5389	10786	10814
Geometric-Waring	3.04	0.49	15174.39	-	-5389	10786	10815
Negative binomial-Waring	3.02	0.49	83.86	0.1	-5389	10787	10823
Discrete Pareto	2.27	-	-	-	-5431	10866	10880
Negative binomial	1.29	0.19	-	-	-5507	11019	11041
Geometric	2.04	-	-	-	-6134	12271	12285
Geometric-Yule	3.12	1903.66	-	-	-6782	13573	13601
Poisson lognormal	1.18	1.72	-	-	-8879	17763	17785
Poisson	1.04	-	-	-	-9573	19150	19164
Household Partnerships							
Females							
Geometric-Waring	2	0.36	214.09	-	-101	210	235
Waring	7.73	1	-	-	-103	211	230
Negative binomial-Waring	2	0.36	361.31	0.2	-101	212	243

Discrete q-exponential	3.44	0.24	-	-	-103	212	231
Discrete Pareto	8.14	-	-	-	-104	212	225
Geometric	1.00	-	-	-	-109	222	235
Negative binomial	3	0.004	-	-	-111	227	237
Males							
Discrete Pareto	6.7	-	-	-	-243	489	502
Geometric	1.01	-	-	-	-243	490	503
Negative binomial	0.17	0.94	-	-	-242	491	510
Waring	43.29	0.99	-	-	-242	491	510
Discrete q-exponential	13.94	2.57	-	-	-242	491	510
Geometric-Waring	34.88	0.99	23.26	-	-242	493	518
Negative binomial-Waring	19.5	0.99	9.68	0.84	-242	495	526

Table S3.2: Proportion of partnership type by age, egocentric data

Age	Live-in, Row % (n/N)	GF/BF, Row % (n/N)	Casual, Row % (n/N)
Males			
15-19	2.1 (58/521)	86.8 (452/521)	11.1 (11/521)
20-24	10 (158/1300)	77.8 (1012/1300)	12.2 (130/1300)
25-29	27.2 (114/1107)	62.5 (692/1107)	10.3 (301/1107)
30-34	49.5 (56/880)	44.1 (388/880)	6.4 (436/880)
35-39	65.9 (49/818)	28.1 (230/818)	6 (539/818)
40-44	78.4 (34/759)	17.1 (130/759)	4.5 (595/759)
45-49	82.7 (20/607)	14 (85/607)	3.3 (502/607)
50-54	84.2 (31/587)	10.6 (62/587)	5.3 (494/587)
55-59	87.4 (19/476)	8.6 (41/476)	4 (416/476)
60+	89.2 (18/706)	8.2 (58/706)	2.5 (630/706)
Females			
15-19	12.4 (13/515)	85 (438/515)	2.5 (64/515)
20-24	25.9 (41/1244)	70.8 (881/1244)	3.3 (322/1244)
25-29	45.3 (42/1275)	51.4 (655/1275)	3.3 (578/1275)
30-34	62.2 (32/1057)	34.8 (368/1057)	3 (657/1057)
35-39	69.6 (30/979)	27.4 (268/979)	3.1 (681/979)
40-44	80.2 (15/839)	18 (151/839)	1.8 (673/839)
45-49	80.7 (9/765)	18.2 (139/765)	1.2 (617/765)
50-54	85.3 (15/617)	12.3 (76/617)	2.4 (526/617)
55-59	94.1 (3/387)	5.2 (20/387)	0.8 (364/387)
60+	93.4 (3/412)	5.8 (24/412)	0.7 (385/412)

Table S3.3: Egocentric Newman's assortativity coefficients with only most recent partner

Characteristic	Assortativity (95% pseudo-pop closed-form CI)	Assortativity (95% pseudo-pop jack-knifed CI)	Assortativity (95% weighted bootstrap CI)
Egocentric Partnerships - Men			
Age	0.922 (0.919-0.925)	0.922 (0.915-0.929)	0.922 (0.915-0.928)
Age among live-in partners	0.902 (0.896-0.908)	0.902 (0.889-0.914)	0.902 (0.892-0.910)
Age among non-live-in girlfriend/boyfriends	0.866 (0.858-0.875)	0.866 (0.847-0.886)	0.866 (0.850-0.882)
Age among casual partners	0.839 (0.803-0.870)	0.839 (0.763-0.916)	0.839 (0.740-0.904)
Age among HIV-negative	0.924 (0.920-0.928)	0.924 (0.916-0.933)	0.924 (0.918-0.932)
Age among HIV-positive	0.829 (0.807-0.849)	0.829 (0.790-0.868)	0.829 (0.796-0.861)
Egocentric Partnerships – Women			
Age	0.911 (0.907-0.915)	0.911 (0.900-0.921)	0.911 (0.904-0.917)
Age among live-in partners	0.899 (0.893-0.905)	0.899 (0.885-0.913)	0.899 (0.890-0.907)
Age among non-live-in girlfriend/boyfriends	0.870 (0.861-0.878)	0.870 (0.842-0.897)	0.870 (0.848-0.890)
Age among casual partners	0.843 (0.791-0.884)	0.843 (0.753-0.934)	0.843 (0.779-0.896)
Age among HIV-negative	0.931 (0.927-0.935)	0.931 (0.923-0.939)	0.931 (0.925-0.937)
Age among HIV-positive	0.795 (0.773-0.814)	0.795 (0.751-0.838)	0.795 (0.760-0.815)

Table S3.4 Description of number of partners reported on degree question by difference between degree and egocentric reports

Degree	**3+ less partners in degree (N=3)**	**2 less partners in degree (N=11)**	**1 less partner in degree (N=88)**	**1 more partner in degree (N=999)**	**2 more partners in degree (N=305)**	**3+ more partners in degree (N=24)**
0	3 (100%)	2 (18%)	71 (81%)	0 (0%)	0 (0%)	0 (0%)
1	0 (0%)	9 (82%)	13 (15%)	562 (56%)	0 (0%)	0 (0%)
2	0 (0%)	0 (0%)	4 (5%)	354 (35%)	39 (13%)	0 (0%)
3+	0 (0%)	0 (0%)	0 (0%)	83 (8%)	266 (87%)	24 (100%)

Figure S3.1: Non-household and household degree distributions for males and females with fitted distributions A) males' non-household partnerships, B) females' non-household partnerships, C) males' household partnerships, D) females' household partnerships

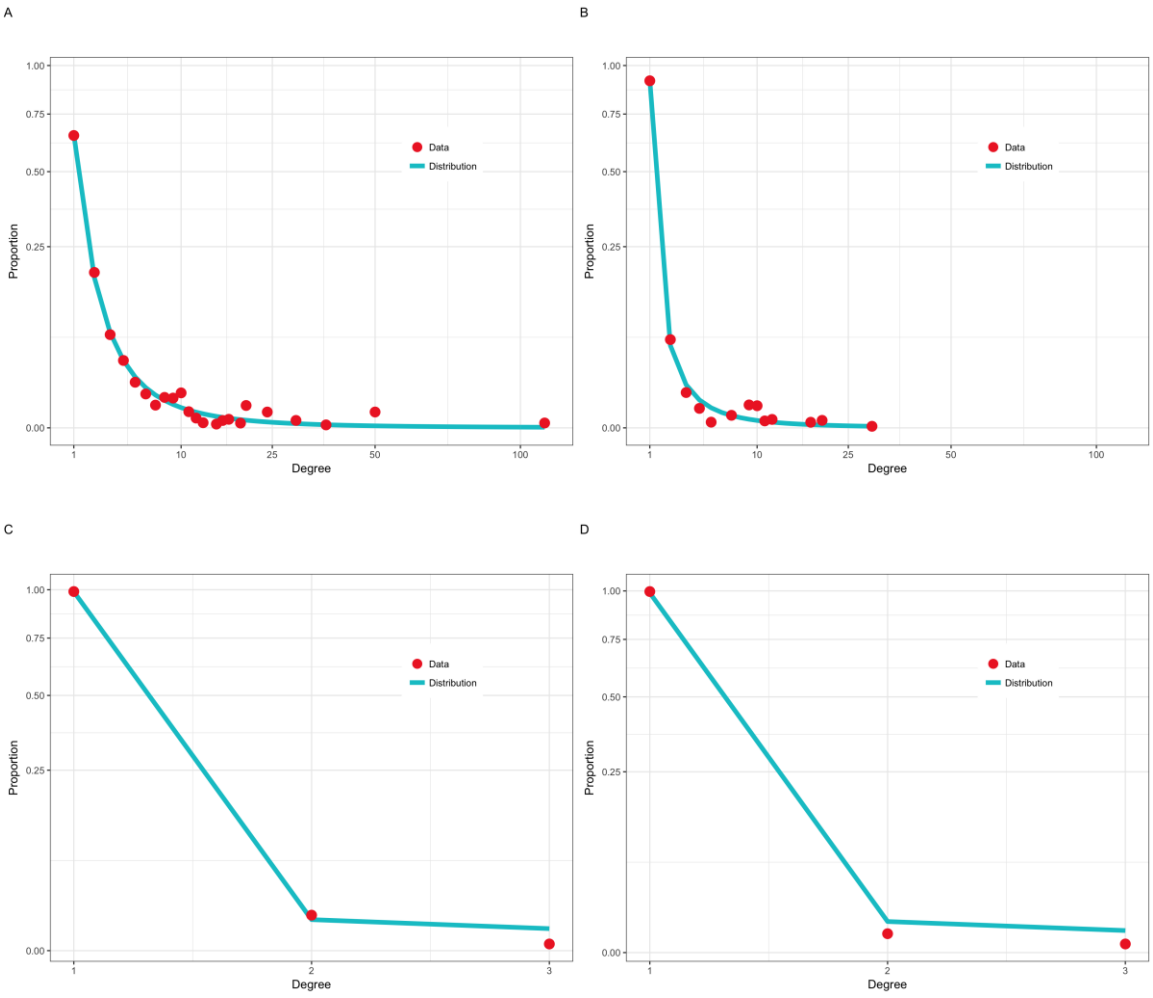


Figure S3.2: Age mixing matrices for egocentric data stratified by HIV status: A) HIV positive male respondents, B) HIV positive female respondents, C) HIV negative male respondents, D) HIV negative female respondents. Each cell represents the relative ratio of observed mixing compared to random mixing.

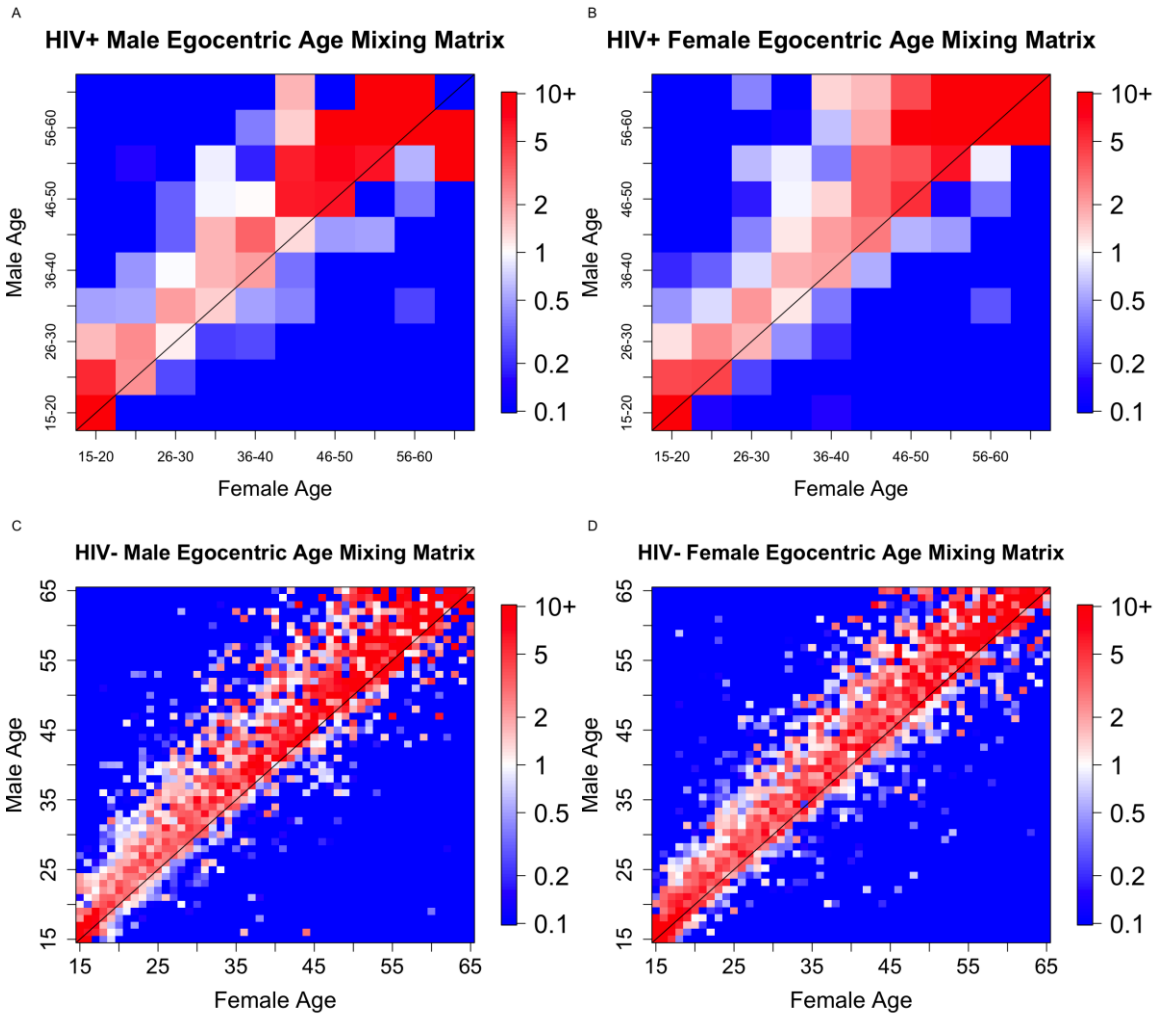
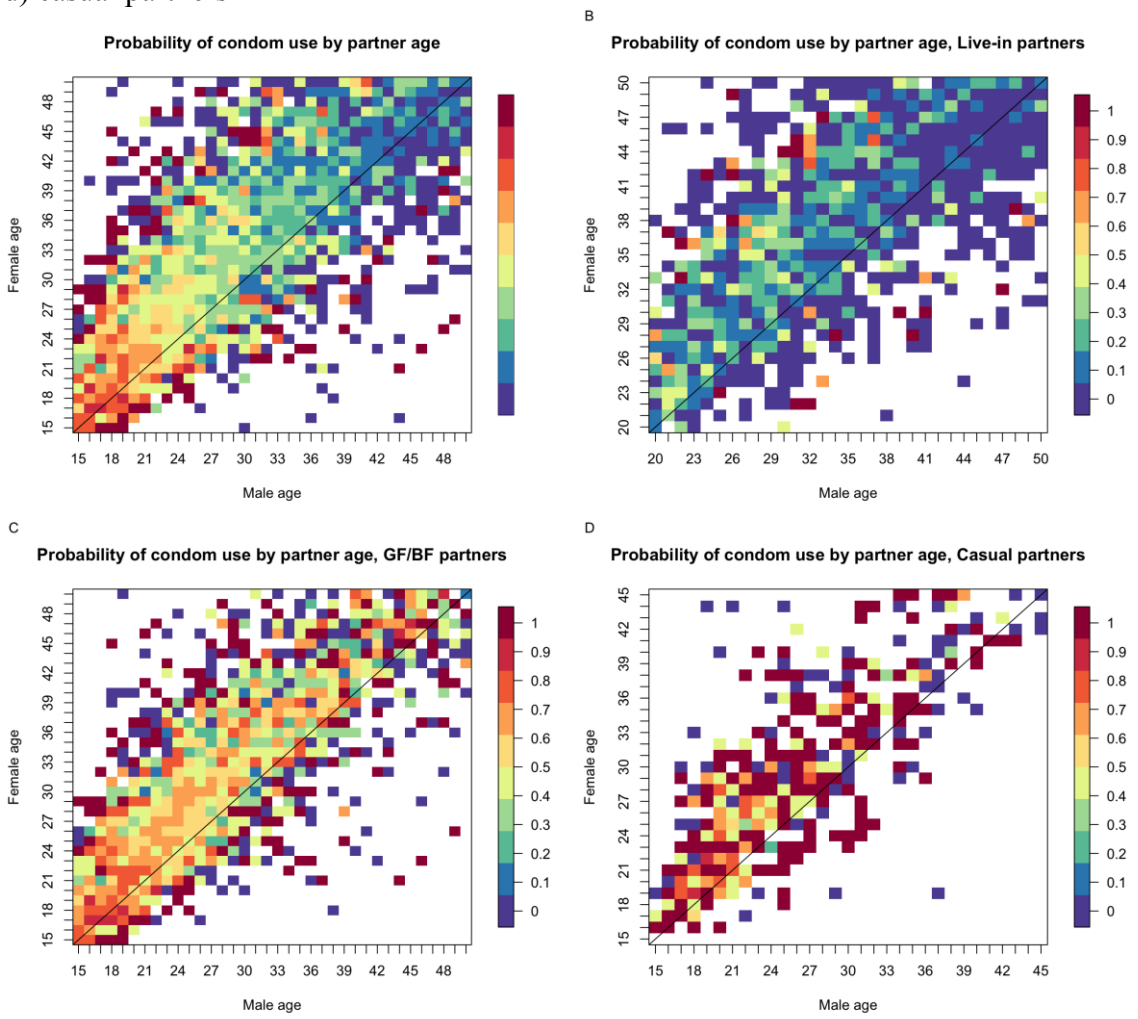


Figure S3.3: Probability of using a condom at last sex by partner age, egocentric data among a) all egocentric data, b) live-in partners, c) non-live-in girlfriends/boyfriends, and d) casual partners



PMFs of fit degree distributions
Poisson

$$P(K = k) = \frac{V_1^k e^{-V_1}}{k!}$$

Poisson with lognormally distributed mean

$$P(K = k) = \frac{e^{kV_1 - e^{V_1}}}{k!} \int_{-\infty}^{\infty} e^{kV_2 u - e^{-V_2 u}} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

Negative binomial distribution

$$P(K = k) = \frac{\Gamma(k + V_1 + V_2)}{\Gamma(V_1 + V_2)k!} * V_2^{V_1 + V_2}(1 - V_2)^k$$

Geometric distribution

$$P(K = k) = \left(1 - \frac{1}{V_1}\right)^{k-1} \frac{1}{V_1}$$

Discrete Pareto distribution

$$P(K = k) = \exp(-V_1 * \log(k) - \log(\zeta(V_1)))$$

where

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

Waring distribution

$$P(K = k|K > 0) = \frac{(V_1 - 1)\Gamma(V_1 + V_2)}{\Gamma(V_2 + 1)} * \frac{\Gamma(k + V_2)}{\Gamma(k + V_2 + V_1)}, V_2 > -1$$

Yule distribution

$$P(K = k) = V_1 \frac{\Gamma(k)\Gamma(V_1 + 1)}{\Gamma(k + V_1 + 1)}, V_1 > 0$$

Discrete Q-exponential

$$P(K = k) = \exp\left(-V_1 * \log\left(1 + \frac{k - 1}{V_2}\right)\right) - \exp\left(-V_1 * \log\left(1 + \frac{k}{V_2}\right)\right)$$

3.7.5. Appendix References

1. Mantel N. The detection of disease clustering and a generalized regression approach. Cancer research. 1967.

4. Simulated sexual networks in South Africa: Implications for HIV transmission

Kathryn A. Risher, David D. Celentano, Shruti H. Mehta, Thomas Rehle, Leickness Simbayi, Olive Shisana, Justin Lessler

4.1. Abstract

Background: Little is understood about sexual networks in South Africa, the country with the world's greatest burden of HIV, despite extensive theoretical understanding of the impact that network structure can play on HIV transmission. Egocentric and household sexual network data is easy to collect and can inform simulations of sexual networks.

Methods: We present a nonparametric approach to simulating sexual networks consistent with incomplete network data using Metropolis-coupled Markov chain Monte Carlo methods. We implement this method using a nationally representative household survey of South Africans, and describe network statistics consistent with limited egocentric and household partnership data. We compare two methods of handling balance between male and female number of partnerships. Finally, we explore how network structure impacts disease transmission on networks consistent with limited data.

Results: Consistent networks are similar across multiple measures of network structure, including number of components, average path length, and average betweenness. Structures are very sensitive to assumptions about structure and reporting of sexual partnerships used to balance inconsistency between numbers reported by male and female participants. Consistent sexual networks in KwaZulu-Natal are less well connected than the rest of the country. Differences in network structure resulting from heterogeneous

assumptions can have a large impact on transmission dynamics, though behavior was similar within consistent networks with shared underlying assumptions.

Conclusions: Simulating sexual networks consistent with limited egocentric and household data allows for improved understanding of disease transmission in South Africa and has potential to impact future HIV intervention implementation.

4.2. Introduction

Sexual contact accounts for the majority of HIV transmission globally[2] and in southern Africa[3]. Previous work has shown that sexual networks can impact the effectiveness of combination HIV prevention interventions[4, 5]. Due to the dyadic (i.e. between two individuals) nature of sexual partnerships, a given individual's risk of HIV acquisition is not purely driven by their own sexual risk behavior, but by the sexual risk behaviors of their sexual partners as well[8]. One can conceive of the sexual partnerships between individuals as the links of a network, and therefore a complete sexual network would describe the individuals in a community and the sexual partnerships between them[9]. A sexual network describes the paths that a sexually transmitted infection, like HIV, is able to take through a community[10]. Thus, in order to fully understand sexual transmission of HIV, it is necessary to understand the complete sexual networks across which HIV spreads[9]. However, sexual networks are extremely difficult to study, partly due to the sensitive nature of sexual behavior, and partly due to difficulties in completely characterizing any network (such as boundary specification – the idea of defining the bounds of individuals included in a network and excluding those that are not[1, 11] - and geographically disparate partnerships)[8, 9]. As such, we have relatively little empirical data on complete sexual networks in sub-Saharan Africa (SSA), the region that has been hardest hit by the HIV pandemic[12].

Network data is traditionally split into two (or more) design types: egocentric and sociometric. Egocentric designs are designs in which you recruit individuals and ask them questions about their network ties. Egocentric data can include information about degree (number of sexual partners) and partnership information such as mixing (who

forms partnerships with whom). In sociometric study designs, you capture the full network by first conducting a census of the area of interest and then asking each individual which other individuals they have ties to (by looking at the census themselves). There are several study designs referred to as partial network designs (such as snowball sampling or contact tracing) which lie along the continuum between egocentric and sociometric designs by sampling individuals and then recruiting from those individuals' networks, and so on[8]. Egocentric study designs have been criticized as allowing for limited inference of sexual network structure[10], though there has been a movement developing to utilize egocentric mixing characteristics to estimate exponential random graph models (ERGMs), a statistical model used to describe networks[13]. Sociometric studies, on the other hand, provide full network structure data, but are exceedingly difficult and sensitive to conduct given that you must conduct a full census, and in an era of vast interconnectedness it is extremely difficult to define a way to limit the bounds of a network.

The most complete empirical data on sexual networks in SSA come from the Likoma Network Study, a sociometric study of sexual networks on Likoma Island, Malawi. The authors described sexual networks in the three years prior to survey and found that the majority (86%) of components were of five individuals or smaller but that these components represented 34% of the respondents, while two-thirds of the population were in 35 (of 256) components of size six or larger[14]. There was a "giant" component of 883 (of 1803) network members, which was characterized by an overall higher average degree (3 vs 1.8) but still had 40% participants who reported 2 or less partners in the three previous years. Notably, Helleringer and Kohler found in this cross-sectional

assessment that the sparser regions of the network had higher HIV prevalence, though they describe that there are demographic and risk factors that likely explain this paradoxical observation[14]. While the study was conducted on an island, they still encountered edge effects, wherein participants reported sexual partnerships with individuals off of the island or with non-spousal partners over 35 years of age, neither of whom were interviewed, and did not include more than 5 most recent partners.

In SSA, network-based research has focused on the role of concurrency, which is multiple sexual partnerships overlapping in time, on HIV epidemics[15-23]. Concurrency is contrasted with serial monogamy, in which each partnership ends before the next partnership begins. Concurrency can be assessed using easily collected egocentric survey data. Much of this emphasis is drawn from the finding that acute HIV (generally defined as the first two months following infection[24, 25]) is a period of greater transmissibility and as such concurrent partnerships would propagate the epidemic by increasing the likelihood of sex acts with other uninfected partners during the acute phase[16]. Arguments against the "concurrency hypothesis" suggest that coital dilution (i.e. lower number of sex acts with higher number of concurrent partners) results in a reduced impact of concurrency on HIV spread[26, 27]. There is mixed evidence for the coital dilution hypothesis[28, 29]. Models of concurrency have shown that even relatively small increases in concurrency (from 56% to 68% concurrent) changes the proportion of the population connected through a sexual network dramatically (from 2% to 64%)[15].

Exponential random graph models (ERGMs) have increasingly been used in mathematical modeling of HIV and STI transmission[4, 30-34], as they provide a statistical model of sexual networks which can be used as the groundwork for a

transmission model. ERGMs estimate the conditional log odds of a tie being added to a network given the current state of its network, and have been extended to allow for the simulation of dynamic networks[35]. Recently, methods to estimate ERGMs from egocentrically sampled network data have been developed and are widely available in statistical packages[36-38]. However, most mathematical models that utilize ERGMs (including the limited number of ERGM-based models of the sub-Saharan epidemic[4, 23, 31, 32]) have not assessed the variability in network realizations consistent with cross-sectional egocentric network data. Authors acknowledge that little is known about the stability of complete network statistics when using egocentric data to estimate ERGM parameters[13].

Within this context, to answer questions about sexual transmission of HIV in South Africa, we must first attempt to make inference around complete sexual networks in South Africa. We use incomplete sexual network data from a South African survey to generate complete sexual networks using a novel, non-parametric approach. We assess the performance of our method to produce sexual networks that are consistent with our limited data inputs, and characterize the extent to which simulated complete networks are similar to one another using network statistics. We compare methods to reconcile men and women's sexual partnership reporting, and compare networks simulated using data from the province most impacted by the HIV epidemic (KwaZulu Natal) to the rest of the country, to describe differences in network structure. Finally, we characterize the impact of network structure on network HIV transmission.

4.3. Methods

4.3.1. South African National HIV Prevalence, Incidence and Behaviour Survey Description

The fourth South African National HIV Prevalence, Incidence and Behaviour Survey (SABSSM IV) was a multi-stage stratified cluster sample of South African households conducted in 2012. SABSSM IV has been described in detail elsewhere[39]. SABSSM IV sampled 1,000 enumeration areas stratified by province, locality type (urban vs rural, formal vs informal), and race[3]. In each enumeration area, 15 households were selected, and within each household, all individuals were invited to participate. Individuals responded to an interviewer-administered questionnaire that included number of sexual partners in the past year and an egocentric partner block on their three most recent sexual partners that included sex, age, and partner type (husband/wife, live-in partner, girlfriend/boyfriend not living with you, casual partner, someone whom you paid for sex, other). Individuals in sexual partnerships with other household members had the study ID of their household partners recorded at the end of their questionnaire, allowing linking full questionnaires of household partners to one another. We did not require that both partners have listed one another but included partnerships where only one partner listed the other.

All youth and adults who participated provided either written or verbal consent, including parent/guardian informed consent for respondents less than 18 years of age.

[3] Although racial discrimination was abolished in 1994, the four racial categories of Black African, White, Coloured and Indian/Asian continue to be used as both social and economic inequalities still prevail as a legacy of the apartheid policies.

SABSSM IV has Institutional Review Board (IRB) approval from the Human Sciences Research Council IRB and the Centers for Disease Control and Prevention Center for Global Health IRB.

4.3.2. Network generation algorithm

Our goal is to simulate a complete network of sexual contacts using a limited set of incomplete sexual partnership data available in SABSSM IV. Because the size of a network increases dramatically with each person added (for instance, a heterosexual network of 3 men and 3 women has 2^{3*3}=512 possible realizations), we are computationally limited on network sizes that are feasible to simulate. We simulate a sexual network of 1,000 individuals under the assumption that the trends we observe are characteristic of larger complete sexual networks in South Africa.

In order to simulate networks consistent with our observed data, we used Metropolis-coupled Markov chain Monte Carlo ($(MC)^3$)[40] to characterize networks of 1,000 individuals (Figure 4.1). We initialized our 1,000 person networks with joint characteristics sampled proportional to SABSSM IV sampling weights. The characteristics assigned to each person were: sex, age, race, employment, and education. Sex was distributed the same in each network (511 females, 489 males), and our network exclusively represents 15-65 year olds. We then randomly distributed the expected number of casual and regular ties in a 1,000-person network between individuals of opposite sex. We define casual partnerships as any non-live-in partner, and regular partnerships as any live-in partner. Starting with this initial network, on each step of the Markov chain Monte Carlo (MCMC), we permuted the network. We tested several different permutation steps (details in the appendix), and ultimately used as our number

of sexual partnerships (y) perturbed at each step: y=1 + Pois(λ=2). Our permutation step involved for each of the y partnerships having a 1/3 probability of adding a tie (½ casual, ½ regular), a 1/3 probability of subtracting a tie (½ casual, ½ regular), and a 1/3 probability of replacing a tie (½ casual, ½ regular). These y changes were made in y randomly sampled partners (except for in version 4 in which 3 changes were all made within the same individual, detailed in appendix). We additionally permuted one person's characteristics (age, race, education and employment) in 10% of permutation steps, replacing the sampled individual's characteristics with characteristics jointly sampled proportional to sampling weights from the data.

For each network, we calculated the log likelihood of the marginal SABSSM IV partnership data having arisen from the simulated network, by generating a probability mass function from the simulated network. This log likelihood function was the sum of the following: males' casual partnership degree distribution (stratified by age <30 and ≥30); males' regular partnership degree distribution (stratified by age <30 and ≥30); females' casual partnership degree distribution (stratified by age <30 and ≥30); females' regular partner degree distribution (stratified by age <30 and ≥30); female's overall degree; male's overall degree; the age difference between partners; the proportion matching by race, education and employment status; and the population's distribution of age, race, education, and employment. See appendix, Table S4.1 and Table S4.2 for data used in network simulations. At each proposed iteration, the simulated network PMF for each of the likelihood components (e.g., the proportion of males <30 years of age with casual degree 0, 1, 2, etc) was multiplied by count observations in our data (e.g., SABSSM IV number of males <30 years of age with casual degree 0, 1, 2, etc) to get the

likelihood of our data having been generated by that network. We compared the log likelihood of the proposed (permuted) network to the network in the prior step of the chain, and if it was greater than the log of a random uniform number (0,1), we accepted the permuted network.

In parallel, each $(MC)^3$ iteration ran 4 MCMC chains, which are run at different "temperatures." The "cold" chain, from which our results come, explores an untransformed likelihood landscape. The three "hot" chains explore a flattened likelihood landscape (transformed by square, cube and fourth roots, respectively). The flattened likelihood landscape allows these chains to avoid being stuck on local maxima. At each iteration of the algorithm, the "cold" likelihood is compared to the three untransformed "hot" likelihoods and if the likelihood of one of the "hot" chains is higher, that chain becomes the "cold" chain.

We started the $(MC)^3$ with 10 random samples and ran for a burn-in period of 1,000,000 iterations. We then ran the $(MC)^3$ chain for 1,000,000 additional iterations and randomly sampled 10 networks from each $(MC)^3$ chain. We assessed convergence of our $(MC)^3$ chains using the Gelman and Rubin R-hat statistic[41], using the cutoff of <1.2 to indicate convergence[42], estimated using the 'asbio' package in R[43].

4.3.3. Network sensitivity analyses

Men and women report an unbalanced number of partnerships across surveys in SSA[44-50]. In our sample, 11,889 (unweighted=10,194) men reported 12,951.4 (unweighted 10,294) partnerships, while 12,043 (unweighted=13,090) women reported 8,492.2 (unweighted 8,908) partnerships. This represents 4,459.2 excess partners reported by men than women. In our first network simulations ("true reporting"), we

make no modifications, and our simulated network degree distributions are in a space between the distributions reported by men and those reported by women. Next, we consider two scenarios in which we assume that male degree distributions are accurate and adjust female degree distributions accordingly. In the first scenario, "female underreporting," we modify our simulated network female casual degree probability mass function with a truncated Poisson distribution, and fit the underreporting Poisson λ for each group (women <30 casual, women ≥30 casual, women overall degree) (see appendix for more details). In our second scenario, "FSW," we assign a small group of women to be "sex workers" and assume a Poisson prior with mean informed by the literature (see appendix for more details).

4.3.4. Large scale network structure

On our simulated networks, we calculated the following network statistics (defined in Box 4.1): number of components, percent of components size 5 of greater, size of largest component, diameter of largest component, average path length, average betweenness, percent of sexually active in components of size 5 or greater, percent of sexually active in the largest component, percent of largest component with 1 partner, percent of those with 1 partner in the largest component, edge density, and average coreness. We use the R package 'igraph' (v1.0.1)[7] to describe network statistics.

4.3.5. Assessing network transmission characteristics

Within our consistent static networks, we simulated HIV transmission in two ways. First ("method A"), we randomly seed each of our networks 10 times with 10 randomly infected individuals, and run our simulation for 10 timesteps, where each

timestep represents 1 year (see appendix for derivation of probability of infection). Each timestep, if an individual is in a partnership with an infected individual, there is a Bernoulli draw to determine if they become infected. We do not model deaths or treatment, as we are attempting to demonstrate solely the impact of network characteristics on epidemic potential, not a realistic disease simulation.

In our second disease simulation, we run our disease simulation for only one timestep, and start each of our 10 disease simulations with from 1-500 infected individuals, and infection status is determined by probability of HIV infection according to SABSSM IV data (as predicted by simulated individuals' sex, age, race, education and employment status, see Table S4.3 for regression model). For the FSW sensitivity analyses, FSWs have probability of HIV infection based on literature (see appendix). The outcome is number of secondary infections that result at 1 timestep into the simulation.

Using our network simulation results, we regress number infected on network structure characteristics to assess if certain network characteristics are associated with greater epidemic potential. For disease simulation method A, our outcome is final number infected and we use mixed effects linear regression (with random intercepts for network to account for the 10 randomly seeded networks' non-independence). For disease simulation method B, our outcome is the average number of secondary infections for each network seeded with 1:500 individuals and we use linear regression with a quadratic term for number of individuals seeded. Interactions between network characteristics and number seeded, and network characteristics and a quadratic term for number seeded, were tested. Due to high collinearity between network characteristics and

interaction terms incorporated in Method B, we use stepwise regression to select our final adjusted models.

All analyses were conducted in R version 3.3.1 (R Foundation for Statistical Computing).

4.4. Results

The network simulation produced networks which converged on most network characteristics (18/20) and which matched our data inputs. See appendix for more details on the performance of the $(MC)^3$ algorithm.

4.4.1. Network structure & variability

Simulated network structure was relatively consistent within a given network generation algorithm, though there were dramatic differences between the true reporting, female underreporting and FSW scenarios (Figure 4.2 and Figure S4.3). In the true reporting algorithm, there were 266-277 components in the simulated networks, with significantly more components in the female underreporting scenario and significantly less in the FSW scenario (see Box 4.1 for network definitions). A minority of components had 5+ individuals in them (2.2-6.6% in true reporting algorithm), but this was significantly higher in the female underreporting scenario and significantly lower in the FSW scenario. The largest component in our true reporting and female underreporting scenarios were relatively small (9-47), but statistically significantly larger in the FSW scenario (138-224). Diameter varied less dramatically, with the true reporting and female underreporting scenarios having diameter ranging from 4-13, though the FSW scenario was significantly higher (8-16). Average path length and average betweenness were also

both significantly higher in the FSW scenario than in the true reporting, with no difference between female underreporting and true reporting for average path length or for average betweenness.

We find that several network characteristics are highly correlated with one another (Figure S4.4). The number of people in the largest component, percent of sexually active in largest component, percent of those with 1 partner in the largest component, average path length, diameter of the largest component, percent edge density, and average betweenness were all positively correlated. The number of components and percent of sexually active population in components size 5+ were negatively correlated.

In comparing networks simulated using KwaZulu-Natal (KZN) to those using the full country data, we found that KZN networks were significantly "lower risk" networks than the networks simulated from all South African data (Figure S4.5). KZN networks had a lower edge density (which matched the lower degree in Table S4.1), a lower number of components, a smaller diameter in the largest component, shorter average path length, and numerous other characteristics that made these networks less well-connected.

4.4.2. Assessment of transmission characteristics

Our HIV transmission simulations using method A (simulation to year 10 on the static networks), show that the true reporting simulation method and female underreporting simulation method yield highly similar trajectories under this transmission setting (Figure 4.3 and Figure S4.6). While there was stochastic variability around transmission for these two network simulation methods (true reporting final number of infections ranged from 12-34, median 17; female underreporting ranged from 11-37, median=17), these trajectories were almost identical. The FSW trajectory, however,

showed dramatically higher growth in number of infections than the other two simulation methods (final number of infections ranged from 12-82, median=26). We find that our KZN trajectories are almost identical to our "true reporting" method with the full South African data (Figure S4.8a).

Our HIV transmission simulations using method B (number infected in the first generation when infecting from 1-500 individuals based on probability of infection based on demographics or FSW status) show a difference between each of the three methods for generating networks (Figure 4.3b, Figure S4.7). The true reporting method results in the fewest secondary infections, followed by the female underreporting scenario, with the most infections observed among the FSW scenario. Secondary infections increase until the network has been seeded with 350-400 infections and then start to decrease as the network is saturated, a phenomenon observed consistently across methods of network generation. Again, we find that our KZN trajectories are quite similar to our "original" method with the full South African data (Figure S4.8b).

4.4.3. Association between network structure and transmission

All network characteristics were significantly associated with our Method A final number of infections in simple linear mixed effects regression (Table S4.4). Increasing number of components and percent of components with size 5 or greater were both associated with a decreased final number of infections. Percent of people in components size 5 or greater, size of largest component, percent of the sexually active population in the largest component, percent of those with 1 partner in the largest component, percent of those in the largest component with 1 partner, diameter of largest component, edge density, and average coreness were all positively associated with increased final number

of infections. Due to high collinearity between parameters, the final model only included two characteristics significantly associated with final number of infections: edge density and average coreness. Average coreness notably changed from a positive association with final number of infections to a negative association.

In modeling our average network secondary infections in Method B, we found that almost all of our interaction terms between network characteristics and number seeded and between network characteristics and the quadratic term for number seeded were statistically significant in simple regression (Table S4.5). The terms for number seeded and number seeded^2 can be thought of as the slope at the intercept, and the shape of the curvature (with negative terms representing an upside down U, positive terms representing a rightside-up U, and terms closer to 0 representing a flatter curve), respectively. The independent effect for these network characteristics (non-interaction terms) were largely not significant, likely because there was very little spread near the intercept. In adjusted regression, percent of components of size 5 or greater, percent of the largest component that had degree 1, and average path length were all associated with increased intercept, while percent of people in components 5 or greater, percent in largest component, and percent of all degree 1's in the largest component were all associated with a decreased intercept. Interaction terms between number seeded and number of components, percent of people in components 5 or greater and percent in the largest component all were associated with a decreased slope at intercept, while diameter of largest component was associated with an increased slope at intercept. Significant interaction terms with the quadratic term included an increase in the quadratic term (equating to a flatter parabola) for number of components, percent of people in

components five or greater, percent in the largest component, and edge density, while a decreased term (equating to a pointier parabola) was found for percent in components size 5 or greater and diameter.

4.5. Discussion

We introduce and assess a nonparametric method for generating sexual networks using limited data, applying this method to the South African HIV epidemic to find that the space of networks consistent with our inputted data is relatively narrow but that there are major ramifications for the choice of method to deal with male-female disparities in degree distribution. Assuming that all excess male partners come from sex workers, as opposed to a diffuse spreading out of underreporting amongst all females, leads to networks that are highly connected and have much greater epidemic potential.

We build on ERGM networks methods for network generation with our novel nonparametric approach. ERGM networks are fit using a similar MCMC sampling process to that described above, however, ERGMs are fit using a parametric ERGM likelihood that has been fit to egocentrically sampled data – thus calculating the likelihood of each tie being added to the network through the ERGM likelihood[13]. First, an ERGM requires the estimation of sufficient statistics from egocentric data of given ERGM parameters of interest (count of a given degree, etc) in a network sampled of the size of the intended ERGM simulation. Next, a network is generated by proposing individual ties and calculating the likelihood of that tie using the estimated ERGM [13]. Our algorithm removes a step from the simulation process, the estimation of sufficient statistics, and simultaneously allows complete flexibility in the development of the likelihood function. For instance, in future iterations, we plan to use full age and race

mixing matrices to generate networks instead of proportion matched. One of the benefits of ERGMs is that they provide a network characterization that is network-size-invariant, i.e. degree does not increase proportional to the size of the network[51]. Our method does the same, by comparing directly to our data degree distribution for each proposed network.

Our proposed method additionally allows for principled modifications to our sampling method, as demonstrated by our "female underreporting" and "FSW" scenarios. "Balancing" degree distributions has been a focus in previous modeling work, particularly in the context of compartmental modeling[52-54]. A previous network simulation study that accounted for balancing by incorporating FSWs into the network similarly found that the incorporation of FSWs into networks led to a small connected core component[55]. Empirical work has suggested that sex workers account for the sex-imbalance in United States surveys[56]. We find that this assumption leads to a much larger size of the largest connected component, and leads to greater epidemic potential. Our female underreporting scenario led to larger numbers of secondary infections but not to a larger number of infections in a static network over time, likely because the structure of the underlying network did not change dramatically from the "true reporting" network that compromised between male and female degree. Our FSW scenarios are different than the previously proposed "core group" scenarios[57-59], some of which explicitly model sex work[60], in that men are not categorized to be "clients" but merely partner with FSWs as needed to fill up their excess casual partnerships. These men could serve as "bridge" populations between FSWs and non-FSW females[61]. Theoretical work[62-64] has supported HIV prevention interventions among key network members, such as FSWs

in our scenario. Our results support the importance of FSW-focused prevention interventions on the structural, behavioral, and biological levels[65].

In sensitivity analyses, our assumption that male degree distributions are completely accurate are likely not completely correct, given that previous work has not ruled out that men overreport to some extent[44-50]. However, this body of work has found disparate results that do not lead to a simple model to correct for female under- and male over-reporting. In our results, we choose to present sensitivity models that assume male reports are accurate. A limitation to our approach is that our two scenarios (one in which all women consistently underreport, and one in which a population of female sex workers is not captured) are likely oversimplifications. It is more likely that, rather, the truth lies somewhere in the middle of these scenarios.

In comparison to the Likoma Network Study (LNS) networks, we find much lower connectivity among our simulated sexual networks[14]. We hypothesize that this is for a variety of reasons. First, the LNS network was over the previous 3 years where our networks were over the last year. Second, the LNS networks only included sexually active individuals. Third, the LNS networks were among individuals aged 18-35, a population more likely to report multiple sexual partners than 15-65. With all of these considerations, we would not rule out that our networks are consistent with their findings but our findings rather are inconclusive.

We find that several network characteristics, most of which are challenging to collect in an egocentric survey, are important predictors of transmission potential. While our methods fairly consistently produce similar networks within a given set of assumptions, the data does not allow us to tease apart between major network differences

(such as the “FSW” vs “female underreporting” scenarios) to truly understand a network’s transmission potential. However, our potential scenarios provide upper and lower bounds for the description of networks that are feasible.

Our finding that sexual networks in KwaZulu-Natal have lower edge density and lower number of individuals in the largest component suggests that KZN is not currently experiencing higher risk sexual networks, despite having the highest HIV prevalence[39] of provinces in South Africa and maintained high HIV incidence[66], particularly among young women[67, 68]. We do not find any differences by network transmission in KZN compared to the full South Africa data. We cannot rule out that there are possible explanations for why our simulated KZN sexual networks are “less risky” than the actual sexual networks in KZN, such as cultural norms around disclosure of personal information to interviewers, or bias in the sample not capturing individuals who report higher number of partners. It is possible that the large HIV epidemic in KZN has led to behavioral change and we are capturing behavior following this change. Our lack of evidence for higher risk sexual networks in KZN suggests that alternate explanations[3], such as the migrant labor, biological co-factors, poverty, or low linkage to care among positives[69] may currently be contributing more significantly to ongoing transmission in KZN. Alternatively, differences in concurrency as compared to serial monogamy within the last year, which are not directly captured in our results, may explain the differences in KZN network risk (60% of KZN respondents who reported multiple partners self-reported concurrency in the past year, while 55% in the rest of South Africa self-reported concurrency).

A major limitation of our method is that we have not yet proposed a method for the change of partnerships over time. Network dynamics strongly impact the spread of infectious disease, and static networks fail to capture these disease dynamics[70, 71]. A second limitation is that we treat all past year partnerships as concurrent partnerships. In our sample, just over half (56%) of those with multiple sexual partners self-reported concurrent partnerships over the past year. Thus, we overestimate the density of our networks perhaps by as much as double. Our static networks fail to capture the timespan of partnerships, and, for instance, give greater importance to "one-off" partnerships than otherwise should be given. This assumption is complicated further by the coital dilution hypothesis, which suggests that the impact of concurrency on disease transmission is reduced by having lower coital frequency with concurrent partnerships[28, 29]. We incorporate lower per year probability of HIV transmission for casual partnerships to partially account for potential coital dilution.

An additional limitation is that we simulate networks that are all of the same size, and do not vary our network population size. Previous work has shown that, unsurprisingly, with varying network size but identical mean degree, the relative size of the largest component, density, and component sizes vary[15]. Smaller networks do not follow asymptotic results for the formation of a "giant component," and instead generate highly connected networks at lower mean degree[15]. An additional limitation is that we chose a slightly higher bound to indicate convergence of our MCMC chains, and not all of our network structure characteristics converge, suggesting that there could be some low-level autocorrelation remaining within our chains that would be eliminated by running our chains for longer and sampling less frequently. Finally, due to computation

limitations, we are unable to run our simulation using larger networks. As such, we are unable to fit estimates for the most likely number of sex workers in our simulated network and instead assign randomly sex worker status at the beginning of iterations. Our results are likely susceptible to variability in the estimate of HIV prevalence among FSW, though a recent study found that HIV prevalence among FSWs ranged from 39.7% in Cape Town to 71.8% in Johannesburg[72], which is not inconsistent with the estimate we utilize. We do not assess the impact of different numbers of sex workers on network structure, only one scenario. Future work will assess whether the size of the sex worker population influences our results or if the presence of any highly-connected women leads to the same results.

Our presented nonparametric network generation method provides a flexible and modifiable means to simulate sexual networks consistent with regularly collected egocentric and household survey data. Models to better characterize the sexual networks of South Africa allow for more refined mathematical models in which to test both basic questions about the nature of the HIV epidemic, as well as assess potential impact of interventions in the context of networks.

Figure 4.1: Schematic of MCMC (one chain of the $(MC)^3$ algorithm)

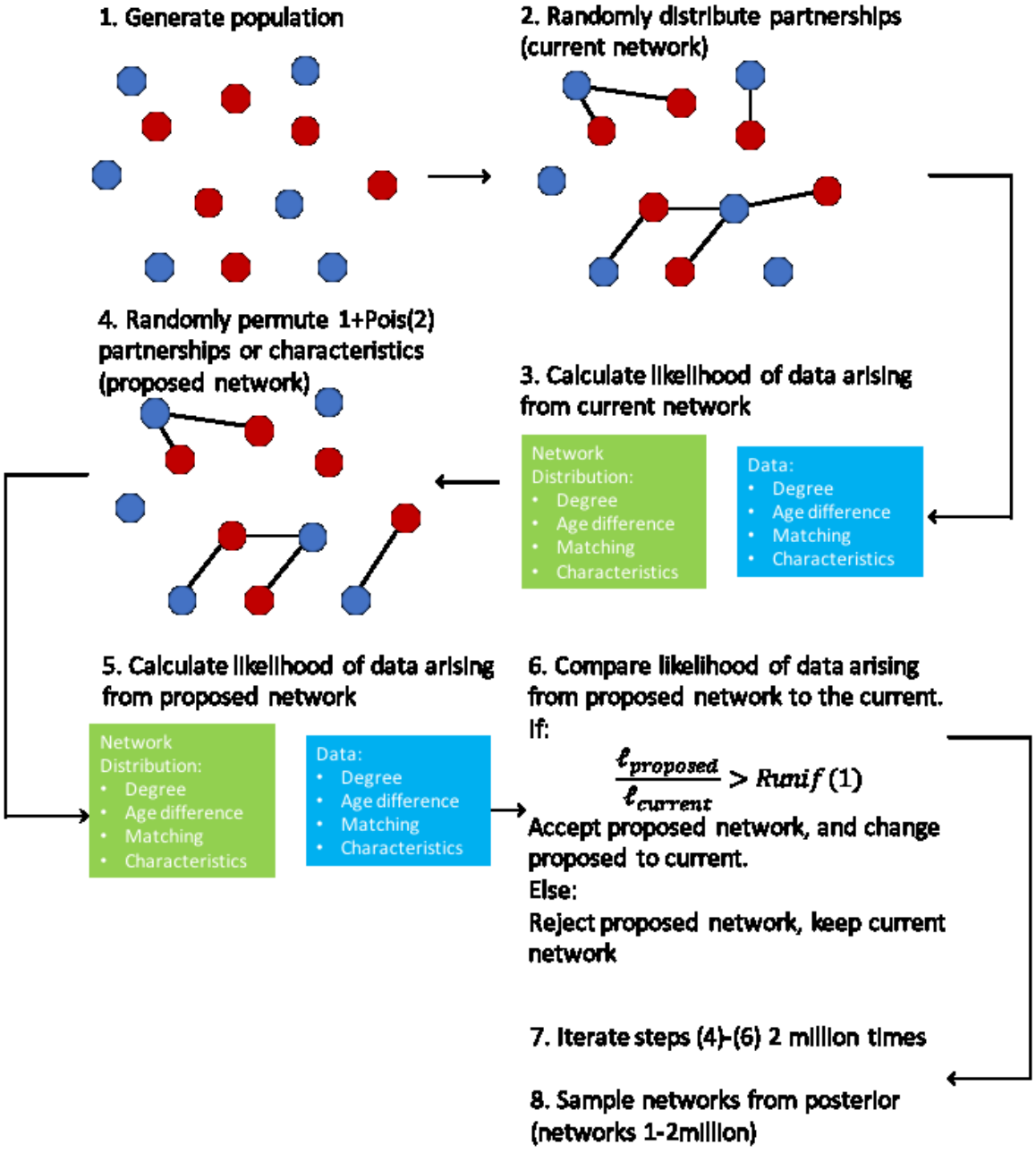


Figure 4.2: Network structure among simulated networks using 1) true reporting algorithm, 2) female underreporting algorithm, 3) missing FSW algorithm. Asterisks (*) indicate a statistically significant difference from true reporting algorithm.

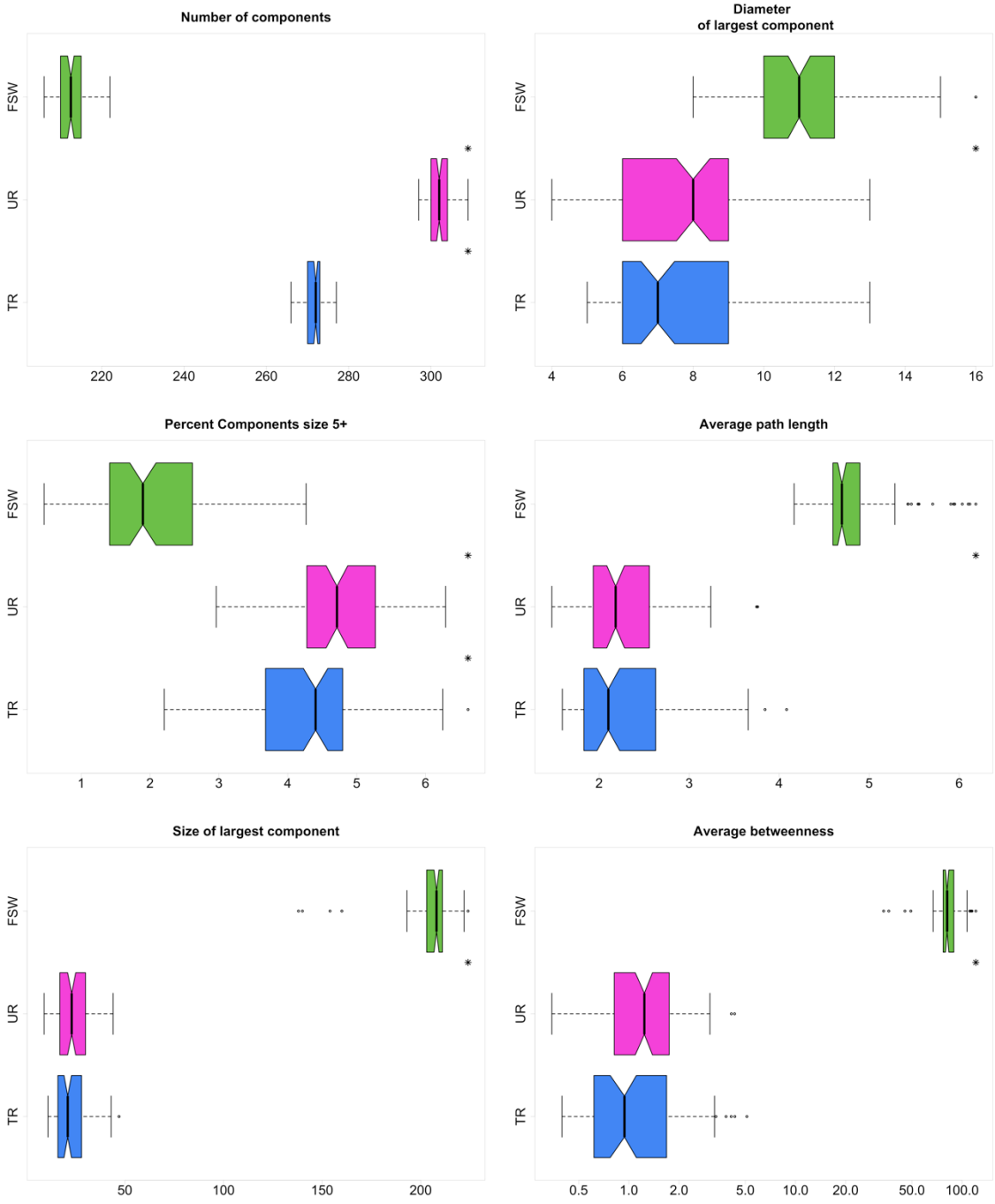
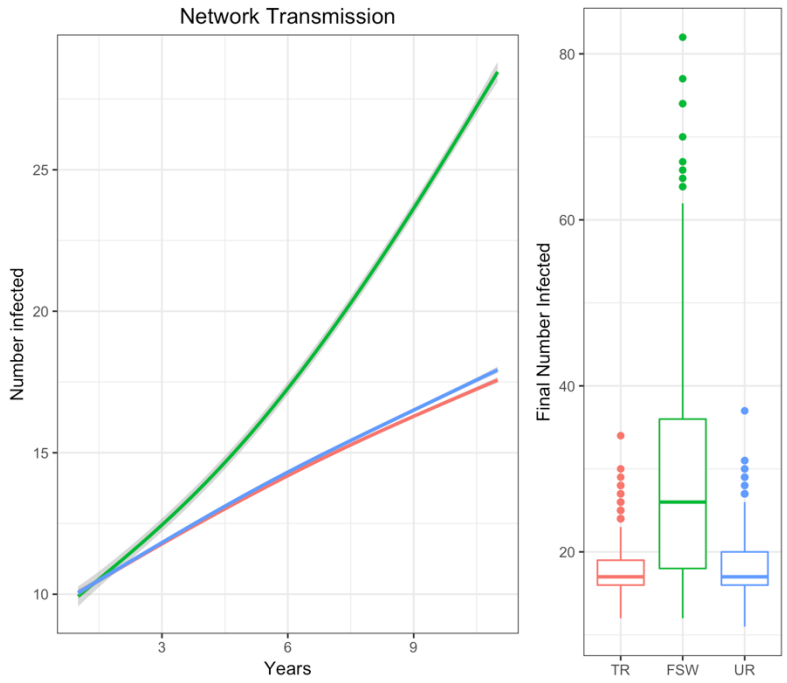
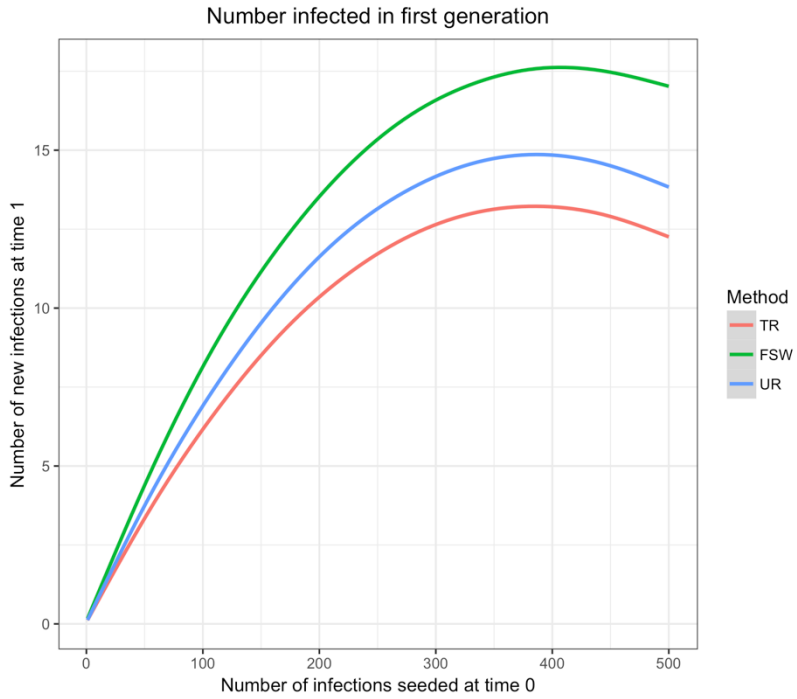


Figure 4.3: Network projections. Panel A: number infected to time-step 10, comparing networks generated with unmodified data (true reporting, TR), underreporting for all females, and missing FSW population. Panel B: number infected in 1 time-step, comparing networks generated with unmodified data (true reporting, TR), underreporting for all females, and missing FSW population.

A



B



Box 4.1: Description of large scale network structure statistics used

- **Components**, subsets of individuals who can be tied to one another through sexual partnerships, but for which there are no sexual partnerships tying to individuals in different subsets[1]
- **Shortest path length** (also referred to in the literature as geodesic distance or distance), the minimum number of partnerships one must step through to get from one individual to another individual on a sexual network[1]
- **Diameter**, the maximum shortest path length between any pair of individuals in a component[1]
- **Edge density**, the number of partnerships observed divided by the number of all possible partnerships[1]
- **Coreness**, using the concept of k-cores, which are maximally connected subsets of a network for which all individuals in the subset have at least degree "k," the "coreness" of a person represents the maximum k-core to which that person belongs[6, 7]
- **Betweenness**, the number of shortest paths that an individual lies along[1]

4.6. References

1. Wasserman S, Faust K. Social Network Analysis: Methods and Applications. New York, NY: Cambridge University Press; 1994.

2. Abdool Karim SS, Abdool Karim Q, Gouws E, Baxter C. Global epidemiology of HIV-AIDS. Infectious disease clinics of North America. 2007;21(1):1-17.

3. Delva W, Abdool Karim Q. The HIV epidemic in Southern Africa - Is an AIDS-free generation possible? Current HIV/AIDS reports. 2014;11(2):99-108.

4. Jenness SM, Goodreau SM, Morris M, Cassels S. Effectiveness of combination packages for HIV-1 prevention in sub-Saharan Africa depends on partnership network structure: a mathematical modelling study. Sexually transmitted infections. 2016;92(8):619-24.

5. Delva W, Helleringer S. Beyond Risk Compensation: Clusters of Antiretroviral Treatment (ART) Users in Sexual Networks Can Modify the Impact of ART on HIV Incidence. PloS one. 2016;11(9):e0163159.

6. Seidman SB. Network structure and minimum degree. Social Networks. 1983;5(3):269-87.

7. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal. 2006;Complex Systems:1695.

8. Morris M, International Union for the Scientific Study of Population. Network epidemiology : a handbook for survey design and data collection. Oxford ; New York: Oxford University Press; 2004. xii, 237 p. p.

9. Delva W, Leventhal GE, Helleringer S. Connecting the dots: network data and models in HIV epidemiology. Aids. 2016;30(13):2009-20.

10. Doherty IA, Padian NS, Marlow C, Aral SO. Determinants and consequences of sexual networks as they affect the spread of sexually transmitted infections. The Journal of infectious diseases. 2005;191 Suppl 1:S42-54.

11. Liljeros F, Edling CR, Nunes Amaral LA. Sexual networks: implications for the transmission of sexually transmitted infections. Microbes Infect. 2003;5(2):189-96.

12. UNAIDS. Global AIDS Update 2016 Available: http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf.

13. Krivitsky PN, Handcock MS, Morris M. Adjusting for Network Size and Composition Effects in Exponential-Family Random Graph Models. Stat Methodol. 2011;8(4):319-39.

14. Helleringer S, Kohler HP. Sexual network structure and the spread of HIV in Africa: evidence from Likoma Island, Malawi. Aids. 2007;21(17):2323-32.

15. Carnegie NB, Morris M. Size matters: concurrency and the epidemic potential of HIV in small networks. PloS one. 2012;7(8):e43048.

16. Eaton JW, Hallett TB, Garnett GP. Concurrent sexual partnerships and primary HIV infection: a critical interaction. AIDS and behavior. 2011;15(4):687-92.

17. Knopf A, Morris M. Lack of association between concurrency and HIV infection: an artifact of study design. Journal of acquired immune deficiency syndromes. 2012;60(1):e20-1; author reply e1.

18. Kretzschmar M, Morris M. Measures of concurrency in networks and the spread of infectious disease. Mathematical biosciences. 1996;133(2):165-95.

19. Maughan-Brown B, Kenyon C, Lurie MN. Partner age differences and concurrency in South Africa: implications for HIV-infection risk among young women. AIDS and behavior. 2014;18(12):2469-76.

20. McCreesh N, O'Brien K, Nsubuga RN, Shafer LA, Bakker R, Seeley J, et al. Exploring the potential impact of a reduction in partnership concurrency on HIV incidence in rural Uganda: a modeling study. Sexually transmitted diseases. 2012;39(6):407-13.

21. Morris M, Epstein H. Role of concurrency in generalised HIV epidemics. Lancet. 2011;378(9806):1843-4; author reply 5-6.

22. Morris M, Kretzschmar M. Concurrent partnerships and the spread of HIV. Aids. 1997;11(5):641-8.

23. 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 and behavior. 2012;16(2):312-22.

24. Bellan SE, Dushoff J, Galvani AP, Meyers LA. Reassessment of HIV-1 acute phase infectivity: accounting for heterogeneity and study design with simulated cohorts. PLoS medicine. 2015;12(3):e1001801.

25. Pilcher CD, Tien HC, Eron JJ, Jr., Vernazza PL, Leu SY, Stewart PW, et al. Brief but efficient: acute HIV infection and the sexual transmission of HIV. The Journal of infectious diseases. 2004;189(10):1785-92.

26. Sawers L. Measuring and modelling concurrency. Journal of the International AIDS Society. 2013;16:17431.

27. Sawers L, Isaac AG, Stillwaggon E. HIV and concurrent sexual partnerships: modelling the role of coital dilution. Journal of the International AIDS Society. 2011;14:44.

28. Gaydosh L, Reniers G, Helleringer S. Partnership concurrency and coital frequency. AIDS and behavior. 2013;17(7):2376-86.

29. Jenness SM, Biney AA, Ampofo WK, Nii-Amoo Dodoo F, Cassels S. Minimal coital dilution in Accra, Ghana. Journal of acquired immune deficiency syndromes. 2015;69(1):85-91.

30. Goodreau SM, Carnegie NB, Vittinghoff E, Lama JR, Fuchs JD, Sanchez J, et al. Can male circumcision have an impact on the HIV epidemic in men who have sex with men? PloS one. 2014;9(7).

31. Khanna AS, Roberts ST, Cassels S, Ying R, John-Stewart G, Goodreau SM, et al. Estimating PMTCT's Impact on Heterosexual HIV Transmission: A Mathematical Modeling Analysis. PloS one. 2015;10(8).

32. Roberts ST, Khanna AS, Barnabas RV, Goodreau SM, Baeten JM, Celum C, et al. Estimating the impact of universal antiretroviral therapy for HIV serodiscordant couples through home HIV testing: insights from mathematical models. Journal of the International AIDS Society. 2016;19(1):20864.

33. Morris M, Kurth AE, Hamilton DT, Moody J, Wakefield S. Concurrent partnerships and HIV prevalence disparities by race: linking science and public health practice. American journal of public health. 2009;99(6):1023-31.

34. Jenness SM, Goodreau SM, Rosenberg E, Beylerian EN, Hoover KW, Smith DK, et al. Impact of the Centers for Disease Control's HIV Preexposure Prophylaxis

Guidelines for Men Who Have Sex With Men in the United States. The Journal of infectious diseases. 2016;214(12):1800-7.

35. Krivitsky PN, Handcock MS. A Separable Model for Dynamic Networks. Journal of the Royal Statistical Society Series B, Statistical methodology. 2014;76(1):29-46.

36. Hunter DR, Handcock MS, Butts CT, Goodreau SM, Morris M. ergm: A Package to Fit, Simulate and Diagnose Exponential-Family Models for Networks. Journal of statistical software. 2008;24(3):nihpa54860.

37. Jenness S, Goodreau SM, Morris M, Beylerian E. Package 'EpiModel'. Package 'EpiModel'. 2015.

38. Krivitsky MPN. Package 'statnet'. Package 'statnet'. 2014.

39. 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: HSRC Press; 2014.

40. Altekar G, Dwarkadas S, Huelsenbeck JP, Ronquist F. Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics. 2004;20(3):407-15.

41. Gelman A, Rubin D. Inference from iterative simulation using multiple sequences. Statistical Science. 1992;7:457-511.

42. Brooks S, Gelman A. General Methods for Monitoring Convergence of Iterative Simulations. Journal of Computational and Graphical Statistics. 1998;7:434-55.

43. Aho K. asbio: A collection of statistical tools for biologists. 2016.

44. 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.

45. Morris M. Telling tails explain the discrepancy in sexual partner reports. Nature. 1993;365(6445):437-40.

46. Wellings K, Collumbien M, Slaymaker E, Singh S, Hodges Z, Patel D, et al. Sexual behaviour in context: a global perspective. Lancet. 2006;368(9548):1706-28.

47. Todd J, Cremin I, McGrath N, Bwanika JB, Wringe A, Marston M, et al. Reported number of sexual partners: comparison of data from four African longitudinal studies. Sexually transmitted infections. 2009;85 Suppl 1:i72-80.

48. Cleland J, Boerma JT, Carael M, Weir SS. Monitoring sexual behaviour in general populations: a synthesis of lessons of the past decade. Sexually transmitted infections. 2004;80 Suppl 2:ii1-7.

49. Mensch BS, Hewett PC, Erulkar AS. The reporting of sensitive behavior by adolescents: a methodological experiment in Kenya. Demography. 2003;40(2):247-68.

50. Clark S, Kabiru C, Zulu E. Do men and women report their sexual partnerships differently? Evidence from Kisumu, Kenya. International perspectives on sexual and reproductive health. 2011;37(4):181-90.

51. Krivitsky PN, Morris M. Inference for social network models from egocentrically-sampled data, with application to understanding persistent racial disparities in HIV prevalence in the US. National Institute for Applied Statistics Research Australia: working paper. 2015.

52. Garnett GP, Anderson RM. Balancing sexual partnerships in an age and activity stratified model of HIV transmission in heterosexual populations. IMA J Math Appl Med Biol. 1994;11(3):161-92.

53. Awad SF, Sgaier SK, Tambatamba BC, Mohamoud YA, Lau FK, Reed JB, et al. Investigating Voluntary Medical Male Circumcision Program Efficiency Gains through Subpopulation Prioritization: Insights from Application to Zambia. PloS one. 2015;10(12):e0145729.

54. Garnett GP, Bowden FJ. Epidemiology and control and curable sexually transmitted diseases: opportunities and problems. Sexually transmitted diseases. 2000;27(10):588-99.

55. Merli MG, Moody J, Mendelsohn J, Gauthier R. Sexual Mixing in Shanghai: Are Heterosexual Contact Patterns Compatible With an HIV/AIDS Epidemic? Demography. 2015;52(3):919-42.

56. Brewer DD, Potterat JJ, Garrett SB, Muth SQ, Roberts JM, Jr., Kasprzyk D, et al. Prostitution and the sex discrepancy in reported number of sexual partners. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(22):12385-8.

57. Stigum H, Falck W, Magnus P. The core group revisited: the effect of partner mixing and migration on the spread of gonorrhea, Chlamydia, and HIV. Mathematical biosciences. 1994;120(1):1-23.

58. Thomas JC, Tucker MJ. The development and use of the concept of a sexually transmitted disease core. The Journal of infectious diseases. 1996;174 Suppl 2:S134-43.

59. Zenilman JM, Ellish N, Fresia A, Glass G. The geography of sexual partnerships in Baltimore: applications of core theory dynamics using a geographic information system. Sexually transmitted diseases. 1999;26(2):75-81.

60. Watts C, Zimmerman C, Foss AM, Hossain M, Cox A, Vickerman P. Remodelling core group theory: the role of sustaining populations in HIV transmission. Sexually transmitted infections. 2010;86 Suppl 3:iii85-92.

61. Morris M, Podhisita C, Wawer MJ, Handcock MS. Bridge populations in the spread of HIV/AIDS in Thailand. Aids. 1996;10(11):1265-71.

62. Albert R, Jeong N, Barabasi AL. Error and attack tolerance of complex networks. Nature. 2001;409:378-82.

63. Dezso Z, Barabasi AL. Halting viruses in scale-free networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2002;65(5 Pt 2):055103.

64. Holme P, Jun Kim B, No Yoon C, Kee Han S. Attack vulnerability of complex networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2002;65:056109.

65. Beyrer C, Crago AL, Bekker LG, Butler J, Shannon K, Kerrigan D, et al. An action agenda for HIV and sex workers. Lancet. 2015;385(9964):287-301.

66. Nel A, Mabude Z, Smit J, Kotze P, Arbuckle D, Wu J, et al. HIV incidence remains high in KwaZulu-Natal, South Africa: evidence from three districts. PloS one. 2012;7(4):e35278.

67. Huerga H, Medicins san Frontieres, Puren A, Bouhenia M, Farhat JB, Welte A, et al., editors. Moderate HIV Incidence and High ART Coverage in Rural KwaZulu-Natal: First Population Based Survey. CROI 2014; 2014; Boston, Massachusetts.

68. Abdool Karim Q, Kharsany AB, Frohlich JA, Werner L, Mlotshwa M, Madlala BT, et al. HIV incidence in young girls in KwaZulu-Natal, South Africa--public health imperative for their inclusion in HIV biomedical intervention trials. AIDS and behavior. 2012;16(7):1870-6.

69. Haber N, Tanser F, Bor J, Naidu K, Mutevedzi T, Herbst K, et al. From HIV infection to therapeutic response: a population-based longitudinal HIV cascade-of-care study in KwaZulu-Natal, South Africa. Lancet HIV. 2017.

70. Danon L, Ford AP, House T, Jewell CP, Keeling MJ, Roberts GO, et al. Networks and the epidemiology of infectious disease. Interdisciplinary perspectives on infectious diseases. 2011;2011:284909.

71. Pastor-Satorras R, Castellano C, Mieghem PV, Vespignani A. Epidemic processes in complex networks. Reviews of modern physics. 2015;87(3):925.

72. UCSF, Anova Health Institute, WHRHI. South African Health Monitoring Study (SAHMS), Final Report: The Integrated Biological and Behavioural Survey among Female Sex Workers, South Africa 2013-2014. San Francisco: UCSF; 2015.

4.7. Appendix

4.7.1. MCMC algorithm

We tested several different permutation steps as follows:

1. One partnership permuted at a time (y=1)
2. One plus a random draw from a Poisson distribution ($\lambda = 2$) number of partners permuted at a time (y=1+pois(λ=2))
3. One plus a random draw from a Poisson distribution ($\lambda = 2$) partnerships permuted at a time in the cold chain (details on cold/hot chains below) with a Poisson draw with a higher λ in the hot chains (25, 25 and 50) (y_{cold}=1+pois(λ=2), $y_{hot1\&2}$=1+pois(λ=25), y_{hot3}=1+pois(λ=50))
4. One partnership permuted at a time 95% of the time, and 5% of the time an individual with 3+ partners had 3 partners permuted (95%: y=1; 5%: y=3 within same individual)

4.7.2. Female sex workers

Data on number of sexual partners reported by female sex workers (FSWs) in sub-Saharan Africa is widely variable. Studies in South Africa have found that FSWs report on average 20 sexual partners per week, though this varies substantially with an IQR of 12-28 partners[1]. However, we know some proportion of weekly partners are regular paying partners or non-paying spouse/boyfriend partners with whom the woman has repeated encounters. In Botswana, FSWs reported 7.6 (6.7-8.5) sexual partners in the past week, and this broke into 5.1 one-time paying customers, 2.3 regular paying partners, 0.7 spouse partners, 0.01 lover/boyfriend, and 0.03 casual/non-paying partners

(it is unclear why these do not sum to 7.6)[2]. Other studies in East Africa have found a median of 1 (IQR=1-2) in the past week[3] and 74.5 in the past 12 months (with 54.8 in the past 6 months and 19.1 in the past 30 days)[4].

Population size estimates for the South African FSW population range from 0.86-0.9%[5, 6]. These estimates are irreconcilable with the number of partnership estimates found in the literature and the excess reported partners by men in our survey. If 0.9% of the female population reported even 5 one-time partners per week (260 partners per year), we would end up with women reporting around 2 times as many partners as men. Thus, we must choose between maintaining the FSW population size and the FSW partner frequency from the literature. Under the assumption that FSW partner frequency data could be biased upwards with increased recruitment of the most active sex workers, and the likelihood that women engage in episodic sex work and thus weekly partner reports might not be maintained throughout the year, we choose the utilize the population size estimate and fit our number of partners for FSW accordingly to men's reports.

4.7.3. Female underreporting scenario

For underreporting among women, our "true" (simulated network) degree distribution is modified using a truncated Poisson distribution as follows:

$$\Pr(obs = 0) = \Pr(True = 0) * \frac{\frac{\lambda^0 e^{-\lambda}}{0!}}{e^{-\lambda}\sum_{i=0}^{0}\frac{\lambda^i}{i!}} + \Pr(True = 1) * \frac{\frac{\lambda^1 e^{-\lambda}}{1!}}{e^{-\lambda}\sum_{i=0}^{1}\frac{\lambda^i}{i!}} +$$

$$\Pr(True = 2) * \frac{\frac{\lambda^2 e^{-\lambda}}{2!}}{e^{-\lambda}\sum_{i=0}^{2}\frac{\lambda^i}{i!}} + \cdots$$

$$\Pr(obs = 1) = \Pr(True = 1) * \frac{\frac{\lambda^0 e^{-\lambda}}{0!}}{e^{-\lambda}\sum_{i=0}^{1}\frac{\lambda^i}{i!}} + \Pr(True = 2) * \frac{\frac{\lambda^1 e^{-\lambda}}{1!}}{e^{-\lambda}\sum_{i=0}^{2}\frac{\lambda^i}{i!}} +$$

$$\Pr(True = 3) * \frac{\frac{\lambda^2 e^{-\lambda}}{2!}}{e^{-\lambda}\sum_{i=0}^{3}\frac{\lambda^i}{i!}} + \cdots$$

$$\Pr(obs = 2) = \Pr(True = 2) * \frac{\frac{\lambda^0 e^{-\lambda}}{0!}}{e^{-\lambda}\sum_{i=0}^{2}\frac{\lambda^i}{i!}} + \Pr(True = 3) * \frac{\frac{\lambda^1 e^{-\lambda}}{1!}}{e^{-\lambda}\sum_{i=0}^{3}\frac{\lambda^i}{i!}} +$$

$$\Pr(True = 4) * \frac{\frac{\lambda^2 e^{-\lambda}}{2!}}{e^{-\lambda}\sum_{i=0}^{4}\frac{\lambda^i}{i!}} + \cdots$$

Where λ represents the mean underreporting among women.

4.7.4. Probability of infection in epidemic simulations

In order to incorporate differences in coital frequency between regular and casual partnerships, we must estimate coital frequency and per-act probability of transmission. Based on Grabowski et al's estimate of per 18-month probability of infection from an HIV seroprevalent household partner of 0.153[7], we back-calculate using the following equation:

$$\Pr(infection) = 1 - (1 - \beta)^{c*t}$$

where c is the coital frequency per unit time, t. We find that even the highest per-act probability of transmission gives us an unreasonable estimate for coital frequency over that 18-month period (0.0008 being the per-act probability of transmission estimated for receptive vaginal intercourse[8], and 207 therefore being the expected number of acts in 18-months). Estimates for number of coital acts per year in Rakai range from 28-96, we assume that it is 90[9]. We therefore, assuming a constant probability of transmission during the 18-month interval, estimate that per act probability of transmission is closer to

$1.23x10^{-3}$. Shisana et al suggest that a non-live in partnership will have ~60% of the coital acts of a live-in partnership[10]. Thus, we estimate per year probabilities of transmission for regular partnerships to be 0.105, and for casual partnerships to be 0.064.

4.7.5. Performance of $(MC)^3$ algorithm

After testing several different proposal mechanisms to improve mixing, we found that the proposal mechanism that most effectively explored the likelihood space was proposing one plus a random draw from a Poisson(λ =2) partnerships at a time. This proposal mechanism achieved an acceptance probability of 0.025 in the cold chains, with acceptance probability in the hot chains 0.061, 0.083 and 0.103. While this acceptance probability is very low, this is due to the complicated multi-dimensional likelihood that we used, and why we utilized such a high number of iterations. Other proposal mechanisms led to lower acceptance probabilities or did not adequately achieve our target statistics. Our chains converged (R-hat<1.2) for 18/20 of our tracked parameters (mean partner age difference; race mixing; education mixing; employment mixing; number of males with 1, 2 and 3+ partners; number of females with 0, 1, 2, and 3+ partners; mean female age; mean male age; number of components; edge density; diameter; average coreness; average betweenness). They failed to converge on number of males with 0 partners (R-hat = 1.41) and size of the largest component (R-hat=1.42). In Figure S4.2, we show network simulated degree and mixing compared to our observed data. In our true reporting algorithm, making no adjustment to female degree, we find that while our simulated degree distributions largely fit the observed data, the $(MC)^3$ algorithm finds a middle ground between the observed degree distributions for women and the observed degree distributions for men. Thus, the simulations have a higher proportion of 0's in

male casual degree distributions than the data, with higher degrees having slightly lower proportions than the data. Vice versa, women's casual degree distributions have a lower proportion of 0's than in the data but a higher proportion of 1's and higher degrees. Mixing (average age difference and proportion matching on race, education, and employment) shows good alignment with the data, with slightly lower proportions of race matching in simulations than observed, likely due to the very high proportion matching in the data (97%) being difficult to achieve in a network of size 1000.

In the female underreporting scenario, we fit 0.65, 0.56 and 0.26 as underreporting Poisson means for casual degree amongst young women, casual degree older women, and overall women's degree distributions.

Table S4.1: Degree data inputs for network MCMC simulation

	Females <30	Females ≥30	Males <30	Males ≥30
South Africa - Overall				
Casual degree				
0	2877 (55.7)	6329 (79.9)	2293 (49.5)	4001 (72)
1	2086 (40.4)	1479 (18.7)	1438 (31)	1095 (19.7)
2	139 (2.7)	92 (1.2)	467 (10.1)	249 (4.5)
3	22 (0.4)	13 (0.2)	183 (3.9)	78 (1.4)
4	10 (0.2)	0 (0)	97 (2.1)	39 (0.7)
5	1 (0)	0 (0)	41 (0.9)	21 (0.4)
6	0 (0)	0 (0)	25 (0.5)	9 (0.2)
7	4 (0.1)	0 (0)	13 (0.3)	1 (0)
8	0 (0)	0 (0)	18 (0.4)	12 (0.2)
9	9 (0.2)	7 (0.1)	5 (0.1)	22 (0.4)
10+*	14 (0.3)	6 (0.1)	57 (1.2)	29 (0.5)
Regular degree				
0	4373 (81.8)	4565 (54.8)	4365 (90.5)	2675 (45.6)
1	969 (18.1)	3758 (45.1)	453 (9.4)	3156 (53.8)
2	6 (0.1)	7 (0.1)	3 (0.1)	33 (0.6)
3	0 (0)	3 (0)	1 (0)	1 (0)
KwaZulu-Natal				
Casual degree				
0	756 (50.3)	1602 (73.8)	550 (46.7)	1018 (63.0)
1	501 (47.1)	273 (25.2)	355 (36.5)	217 (28.8)
2	26 (2.5)	8 (2.9)	83 (8.9)	37 (3.1)
3	1 (0.1)	2 (0.1)	32 (3.8)	16 (2.5)
4	0 (0)	1 (0)	13 (1.7)	4 (0.2)
5	0 (0)	0 (0)	5 (1.0)	2 (0.7)
6	0 (0)	0 (0)	2 (0.9)	1 (0.3)
7	0 (0)	0 (0)	3 (0.1)	1 (0.1)
8	0 (0)	0 (0)	2 (0.1)	0 (0)
9	0 (0)	3 (0.3)	1 (0)	11 (0.8)
10+*	1 (0.1)	3 (0.3)	2 (0.2)	7 (0.5)
Regular degree				
0	1146 (91.4)	1084 (64.4)	994 (93.6)	527 (56.7)
1	186 (8.6)	868 (35.6)	83 (6.1)	808 (43.3)
2	0 (0)	1 (0)	2 (0.3)	1 (0)
3	0 (0)	0 (0)	0 (0)	0 (0)

*Full casual degree data is utilized in our MCMC input, but due to small numbers, we condense for presentation purposes

Table S4.2: Additional data inputs for network MCMC simulation and parameters for epidemic simulations

	Value
Partnership characteristics	
Age difference (Male-Female)	3.67 years
% Race Matching	97.0%
% Education Matching	75.5%
% Employment Matching	53.5%
Individual characteristics	
Age* Mean (Median)	34.0 (32)
Race	
Black	78.6%
Coloured	9.3%
Indian/Asian	2.7%
White	9.3%
Education	
Primary	29.1%
Secondary	56.0%
Tertiary	14.8%
Employment Status	
Employed	39.2%
Unemployed	60.8%
Probability of HIV Transmission	
Regular (household) partnerships	0.1047955 [7, 9, 10]
Casual (non-household) partnerships	0.06426394 [7, 9, 10]
HIV Prevalence among FSW	59.6% [11]

*All age data went into MCMC, mean/median presented here for ease of presentation

Table S4.3: Weighted logistic regression model predicting HIV status used in simulations

Variable	aOR	95% CI
Sex		
Female	Ref	Ref
Male	0.19	(0.08-0.44)
Education		
Primary	Ref	Ref
Secondary	0.98	(0.78-1.22)
Tertiary	0.33	(0.2-0.56)
Age		
15-20	Ref	Ref
21-25	5.03	(3.41-7.44)
26-30	8.59	(5.69-12.95)
31-35	12.78	(8.53-19.14)
36-40	9.58	(6.38-14.38)
41-45	6.99	(4.45-10.98)
46-50	4.61	(3.03-7.04)
51-55	3.23	(2.07-5.05)
56-60	1.36	(0.79-2.32)
61-65	0.88	(0.51-1.52)
Race		
Black	Ref	Ref
Coloured	0.13	(0.09-0.19)
Indian/Asian	0.03	(0.01-0.07)
White	0.02	(0-0.12)
Employment		
Employed	Ref	Ref
Unemployed	1.17	(0.91-1.5)
Interactions with Male Sex		
Male*Education		
Primary	Ref	Ref
Secondary	0.97	(0.66-1.42)
Tertiary	0.85	(0.36-2)
Male*Age		
15-20	Ref	Ref
21-25	2.03	(0.91-4.52)
26-30	3.01	(1.32-6.84)
31-35	3.11	(1.37-7.05)
36-40	5	(2.16-11.6)
41-45	3.66	(1.52-8.8)
46-50	4.38	(1.84-10.44)
51-55	5.98	(2.42-14.79)
56-60	5.53	(1.98-15.48)
61-65	10.98	(3.98-30.32)
Male*Race		
Black	Ref	Ref
Coloured	1.38	(0.74-2.59)
Indian/Asian	1.16	(0.32-4.24)
White	0.16	(0.01-2.6)
Male*Employment		
Employed	Ref	Ref
Unemployed	0.68	(0.46-1.01)

Table S4.4: Association of network structure characteristics with final epidemic size. Linear mixed effects model regressions for final epidemic size for Method A.

Network characteristic	Unadjusted coefficient (95%CI)	Adjusted coefficient (95% CI)
Number of components	-0.22 (-0.24, -0.21)	-
Percent of components size 5+	-6.17 (-6.65, -5.68)	-0.40 (-1.03, 0.23)
Percent of people in components size 5+	0.86 (0.82, 0.90)	-
Size of largest component	0.09 (0.09, 0.10)	-
Percent of sexually active in largest component	0.63 (0.60, 0.66)	-
Percent of those with 1 partner in largest component	0.75 (0.72, 0.79)	-
Percent of those in the largest component with 1 partner	0.89 (0.73, 1.04)	-
Diameter of largest component	3.06 (2.49, 3.62)	-
Edge Density	961 (887, 1035)	1240 (1121, 1359)
Average coreness	0.39 (0.02, 0.75)	-1.60 (-1.82, -1.37)

Table S4.5: Association of network structure characteristics with average number of secondary infections with increasing number of seeded infected individuals (1-500) – Method B. Linear regression models with quadratic terms for number of infected individuals seeded.

Network Characteristic	# Seeded	# Seeded^2	Net. Char.	Net. Char * # Seeded	Net. Char * # Seeded^2
Unadjusted Linear Regression Coefficients					
None	0.077 (0.077,0.077)	-1e-04 (-1e-04,-1e-04)	-	-	-
# comp	0.1 (0.099,0.11)	-0.00017 (-0.00018, -0.00016)	-0.0049 (-0.0065, -0.0033)	-9.3e-05 (-0.00011, -7.9e-05)	2.4e-07 (2.2e-07,2.7e-07)
% comp 5+	0.079 (0.077,0.08)	-0.00011 (-0.00011, -1e-04)	-0.0019 (-0.028,0.024)	-0.00034 (-0.00058, -9.8e-05)	7e-07 (2.4e-07,1.2e-06)
% ppl in comp 5+	0.078 (0.078,0.079)	-1e-04 (-0.00011, -1e-04)	-0.0026 (-0.0053, 1e-04)	-6.3e-05 (-8.8e-05, -3.8e-05)	1.3e-07 (8.6e-08,1.8e-07)
Lgst Comp N	0.078 (0.077,0.079)	-1e-04 (-1e-04, -1e-04)	-0.0091 (-0.028, 0.0094)	-0.00027 (-0.00044, -1e-04)	5e-07 (1.7e-07,8.3e-07)
% lgst comp	0.078 (0.078,0.079)	-1e-04 (-0.00011, -1e-04)	-0.017 (-0.045,0.012)	-5e-04 (-0.00076, -0.00024)	9.8e-07 (4.7e-07,1.5e-06)
% deg 1 lgst	0.078 (0.076,0.08)	-1e-04 (-0.00011, -1e-04)	2.1e-05 (-0.0032, 0.0033)	-1.6e-05 (-4.6e-05, 1.4e-05)	3.1e-08 (-2.7e-08,8.8e-08)
% all deg 1's lgst	0.078 (0.077,0.079)	-1e-04 (-1e-04, -1e-04)	-0.0047 (-0.018, 0.0087)	-6.9e-05 (-0.00019, 5.5e-05)	1.3e-07 (-1.1e-07,3.7e-07)
Diameter	0.099 (0.095,0.1)	-0.00016 (-0.00016, -0.00015)	-14 (-18,-9.6)	-0.26 (-0.3, -0.22)	0.00069 (0.00061,0.00077)
Edge density	0.079 (0.078,0.08)	-1e-04 (-0.00011, -1e-04)	-0.03 (-0.074, 0.014)	-0.00067 (-0.0011, -0.00026)	1.4e-06 (5.8e-07,2.1e-06)
Avg path l	0.1 (0.097,0.1)	-0.00016 (-0.00017, -0.00016)	-1.8 (-2.4, -1.2)	-0.035 (-0.04,-0.029)	9e-05 (8e-05,1e-04)
Avg coreness	0.078 (0.077,0.078)	-1e-04 (-1e-04, -1e-04)	-0.023 (-0.05, 0.0037)	-0.00054 (-0.00078, -0.00029)	1.1e-06 (6.7e-07,1.6e-06)
Avg betw	0.082 (0.08,0.084)	-0.00011 (-0.00012, -0.00011)	-0.0077 (-0.022, 0.0064)	-0.00029 (-0.00042, -0.00016)	6.3e-07 (3.8e-07,8.9e-07)

Network Characteristic	Net. Char.	Net. Char * # Seeded	Net. Char * # Seeded^2
Adjusted Linear Regression Coefficients			
None (# Seeded)	-	0.11 (0.1,0.11)	-0.00017 (-0.00019, -0.00016)
# comp	-0.0095 (-0.029,0.01)	-0.00025 (-0.00043, -6.7e-05)	4.8e-07 (1.3e-07, 8.3e-07)
% comp 5+	0.11 (0.073,0.14)	-	-7.3e-07 (-1e-06,

			-4.3e-07)
% ppl in comp 5+	-0.043 (-0.068,-0.017)	-0.00027 (-0.00045, -9.2e-05)	7e-07 (3.3e-07, 1.1e-06)
Lgst Comp N	-	-	-
% lgst comp	-0.26 (-0.34,-0.17)	-0.00091 (-0.0012, -0.00057)	1.6e-06 (9.5e-07, 2.3e-06)
% deg 1 lgst	0.31 (0.22,0.4)	-	-
% all deg 1's lgst	-0.0089 (-0.013, -0.0053)	-6.1e-06 (-1.5e-05, 2.3e-06)	-
Diameter	0.0067 (-0.016,0.03)	0.00035 (0.00015,0.00055)	-8.2e-07 (-1.2e-06, -4.4e-07)
Edge density	-38 (-120,44)	-0.71 (-1.4,0.035)	0.0015 (2.3e-05,0.0029)
Avg path l	0.14 (0.05,0.22)	-	-
Avg coreness	6 (-10,22)	0.15 (-0.0027,0.3)	-
Avg betw	-	-	-0.00027 (-0.00056, 2e-05)

Figure S4.1: Example of plotted simulated network (sexually active population), true reporting

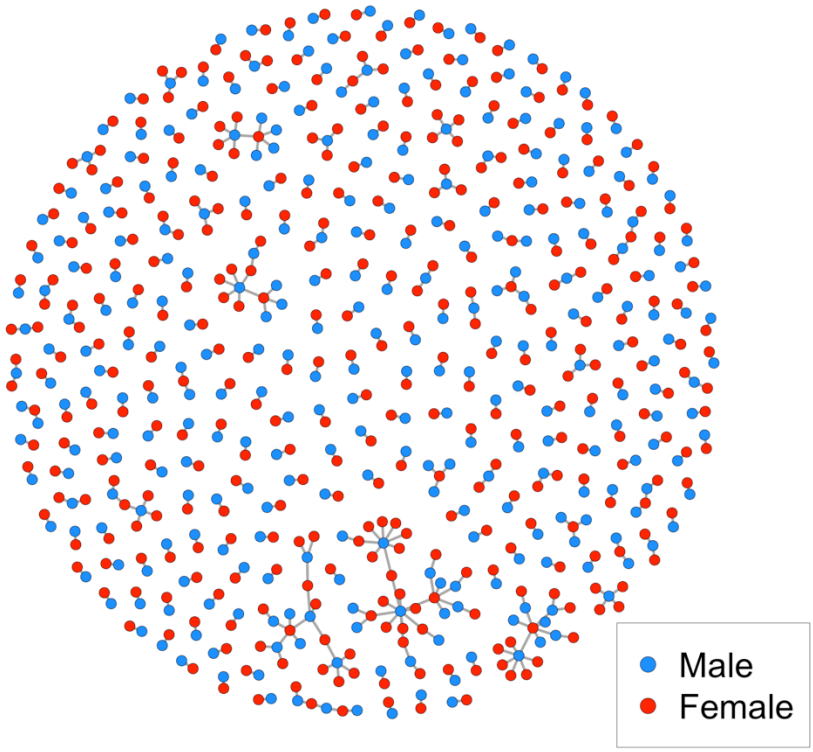
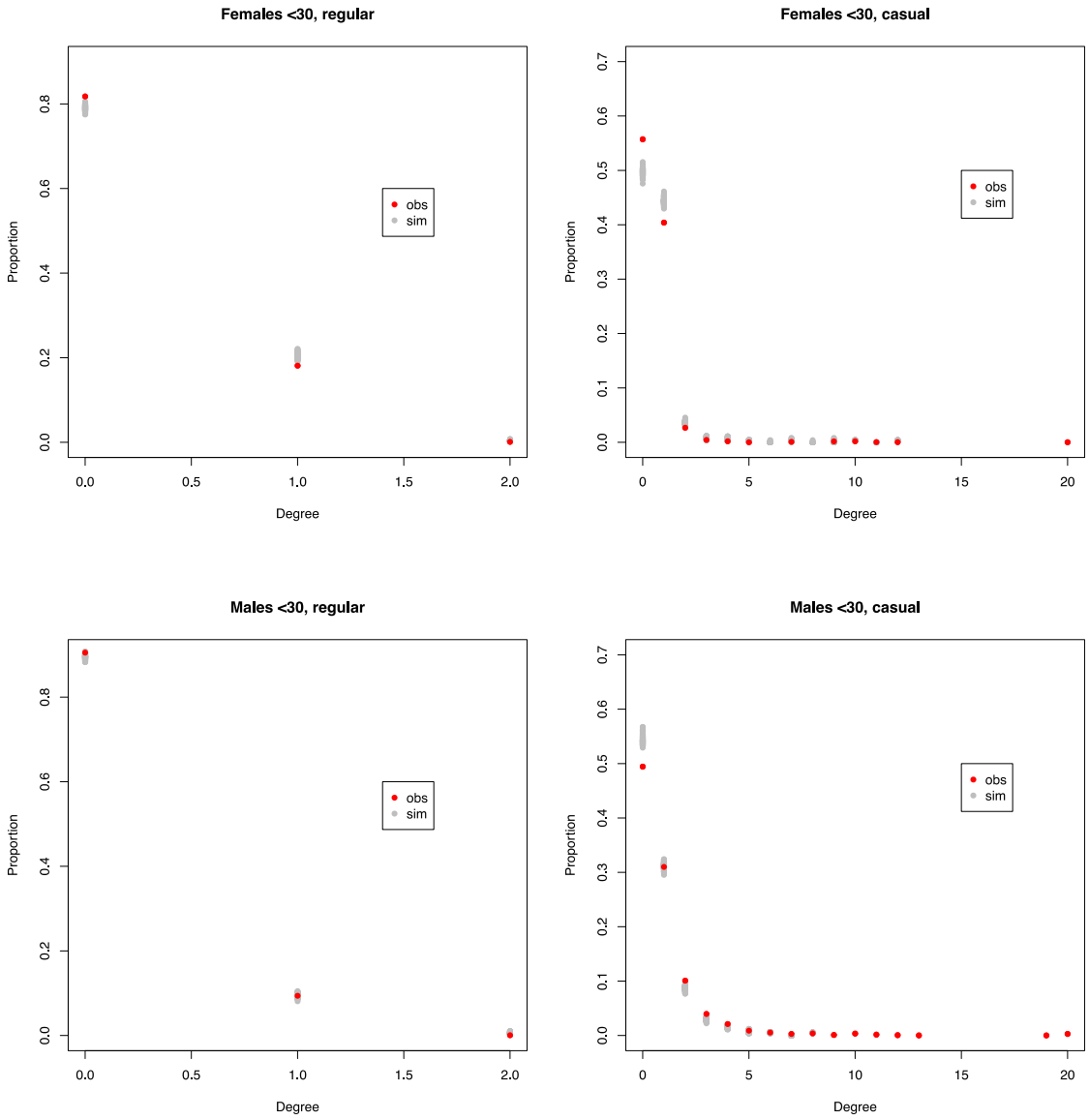
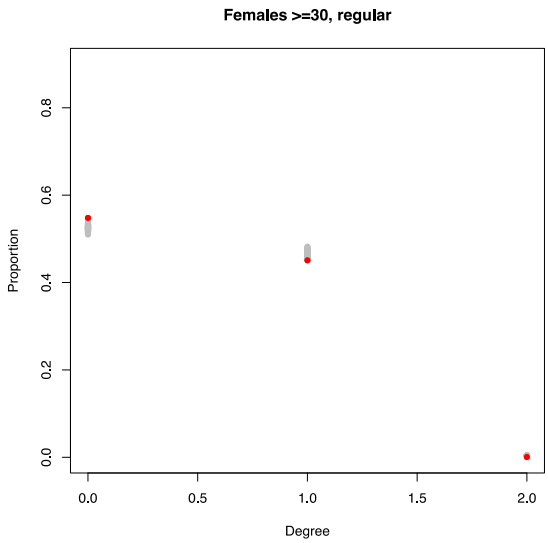
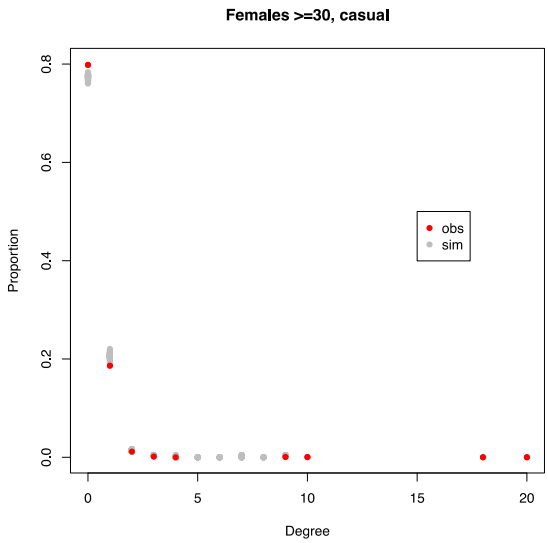
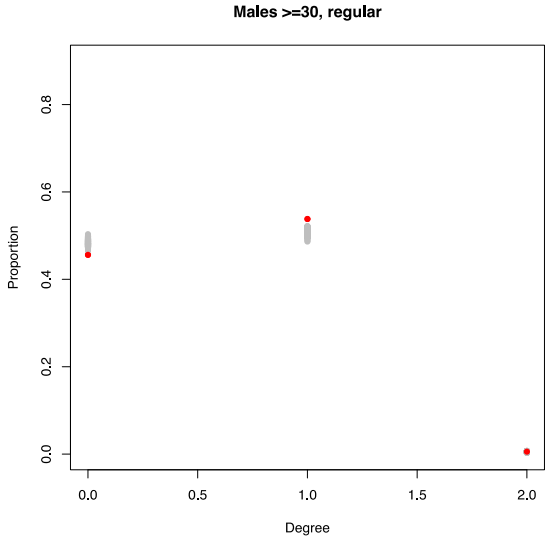


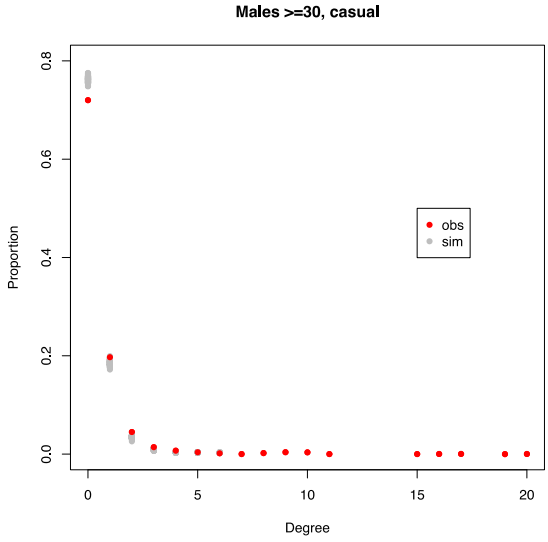
Figure S4.2: Fitted degree distribution compared to data for true reporting algorithm




Females >=30, regular
Proportion
Degree
0.0
0.2
0.4
0.6
0.8
0.0
0.5
1.0
1.5
2.0


Females >=30, casual
Proportion
Degree
obs
sim
0.0
0.2
0.4
0.6
0.8
0
5
10
15
20


Males >=30, regular
Proportion
Degree
0.0
0.2
0.4
0.6
0.8
0.0
0.5
1.0
1.5
2.0


Males >=30, casual
Proportion
Degree
obs
sim
0.0
0.2
0.4
0.6
0.8
0
5
10
15
20

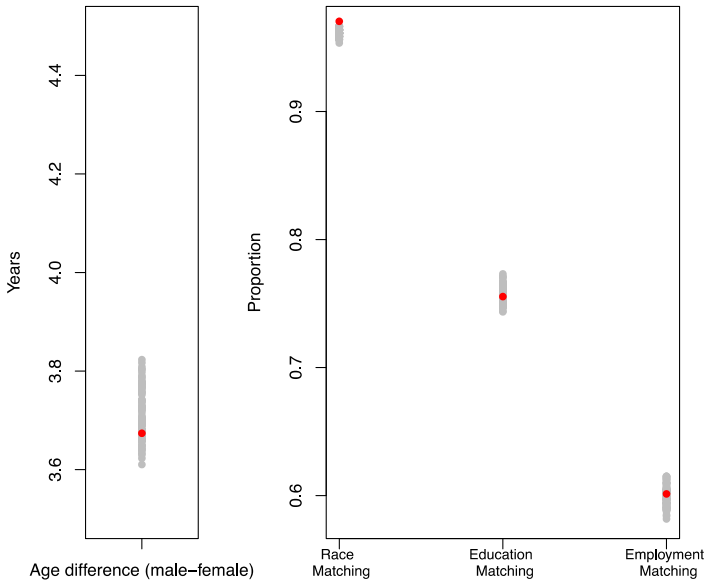

Years
3.6
3.8
4.0
4.2
4.4
Age difference (male−female)
Proportion
0.6
0.7
0.8
0.9
Race Matching
Education Matching
Employment Matching

Figure S4.3: Further description of simulated networks. Asterisks (*) indicate a statistically significant difference from true reporting algorithm.

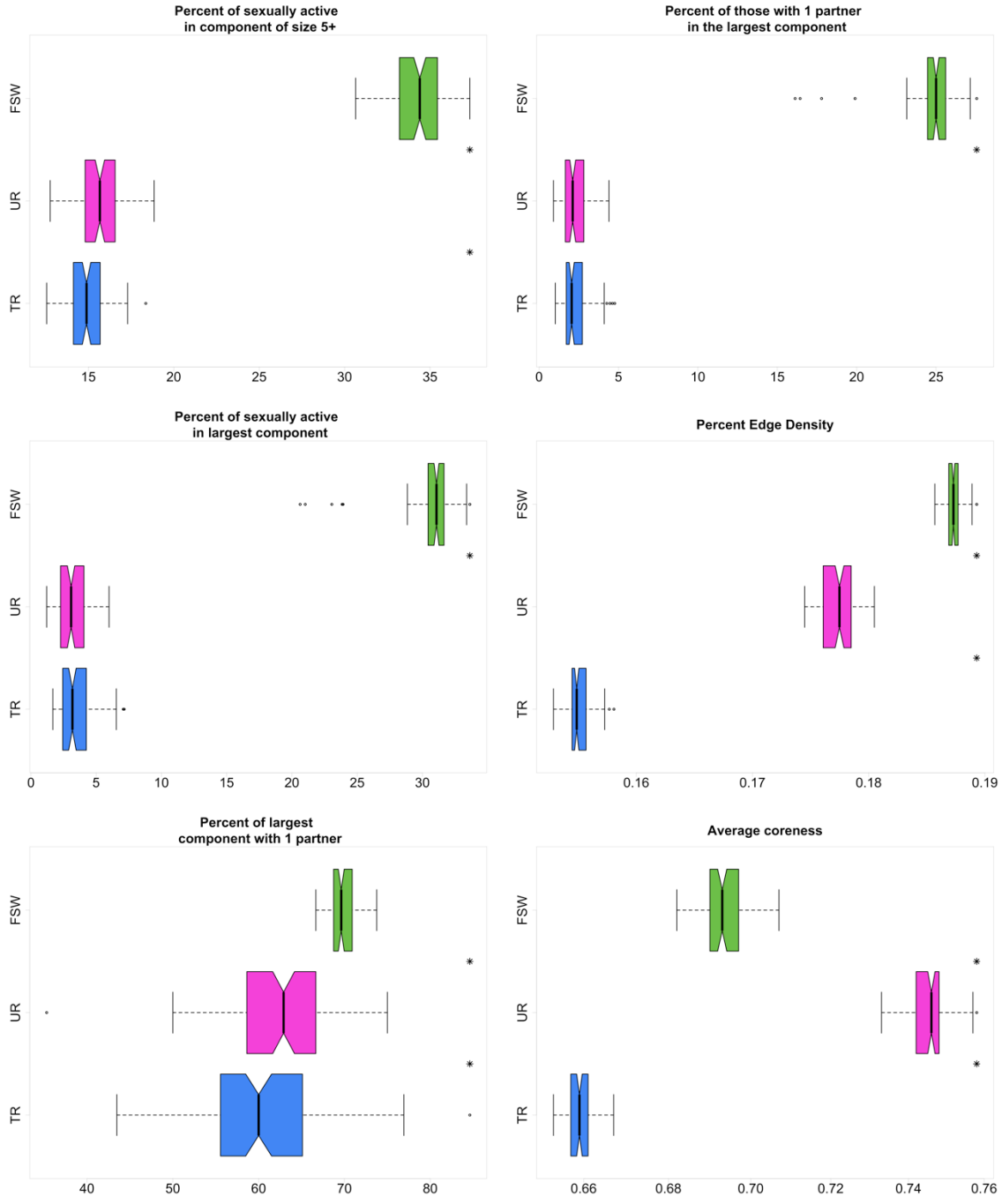


Figure S4.4: Pairwise correlation between network structure statistics in “true reporting” networks. a) Number of components, b) percent of components size 5+, c) percent of sexually active population in components size 5+, d) number of people in largest component, e) percent of sexually active in the largest component, f) percent of those with 1 partner in the largest component, g) percent of largest component with 1 partner, h) diameter of the largest component, i) percent edge density, j) average path length, k) percent average coreness, l) average betweenness.

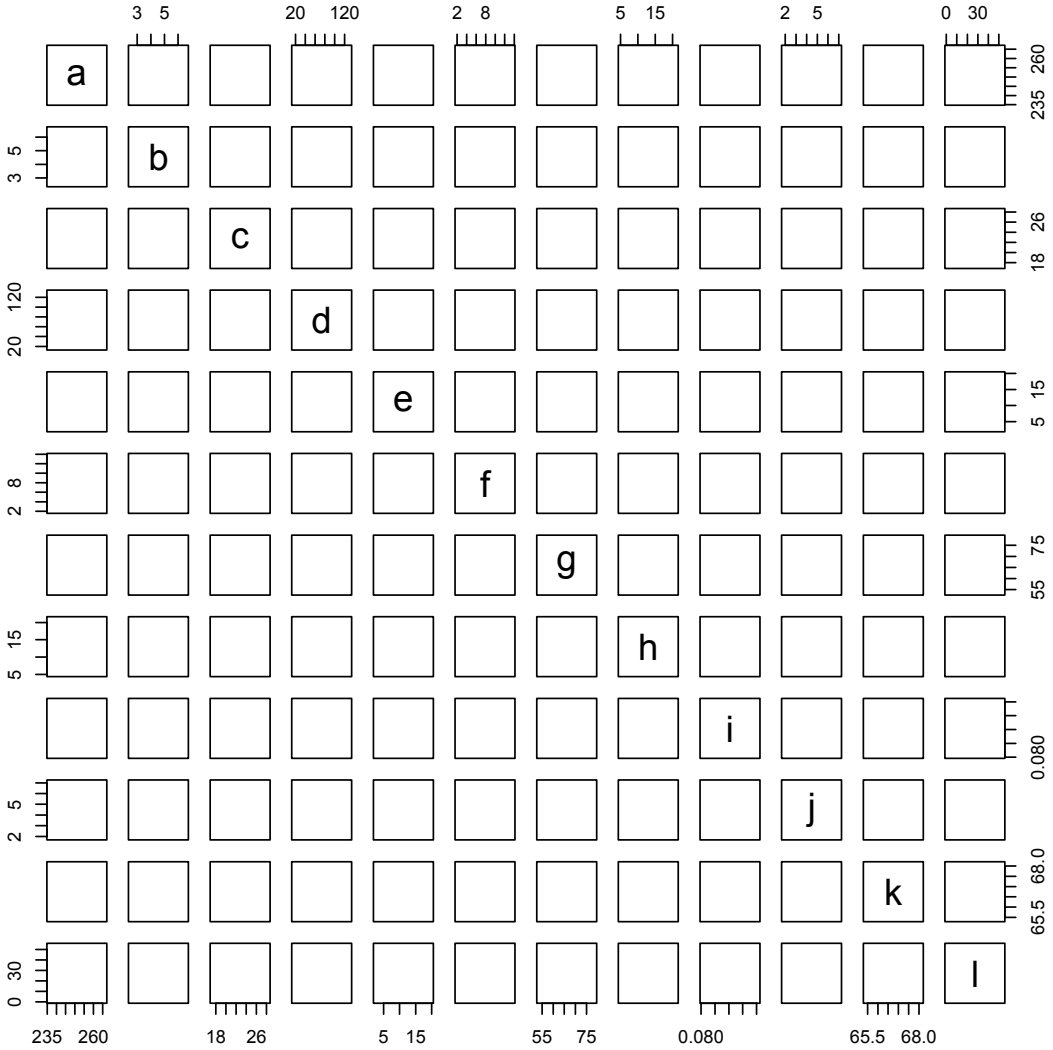


Figure S4.5: Network structure comparing networks fitted to KwaZulu-Natal data with all of South Africa data. Asterisks (*) indicate a statistically significant difference between All RSA and KZN.

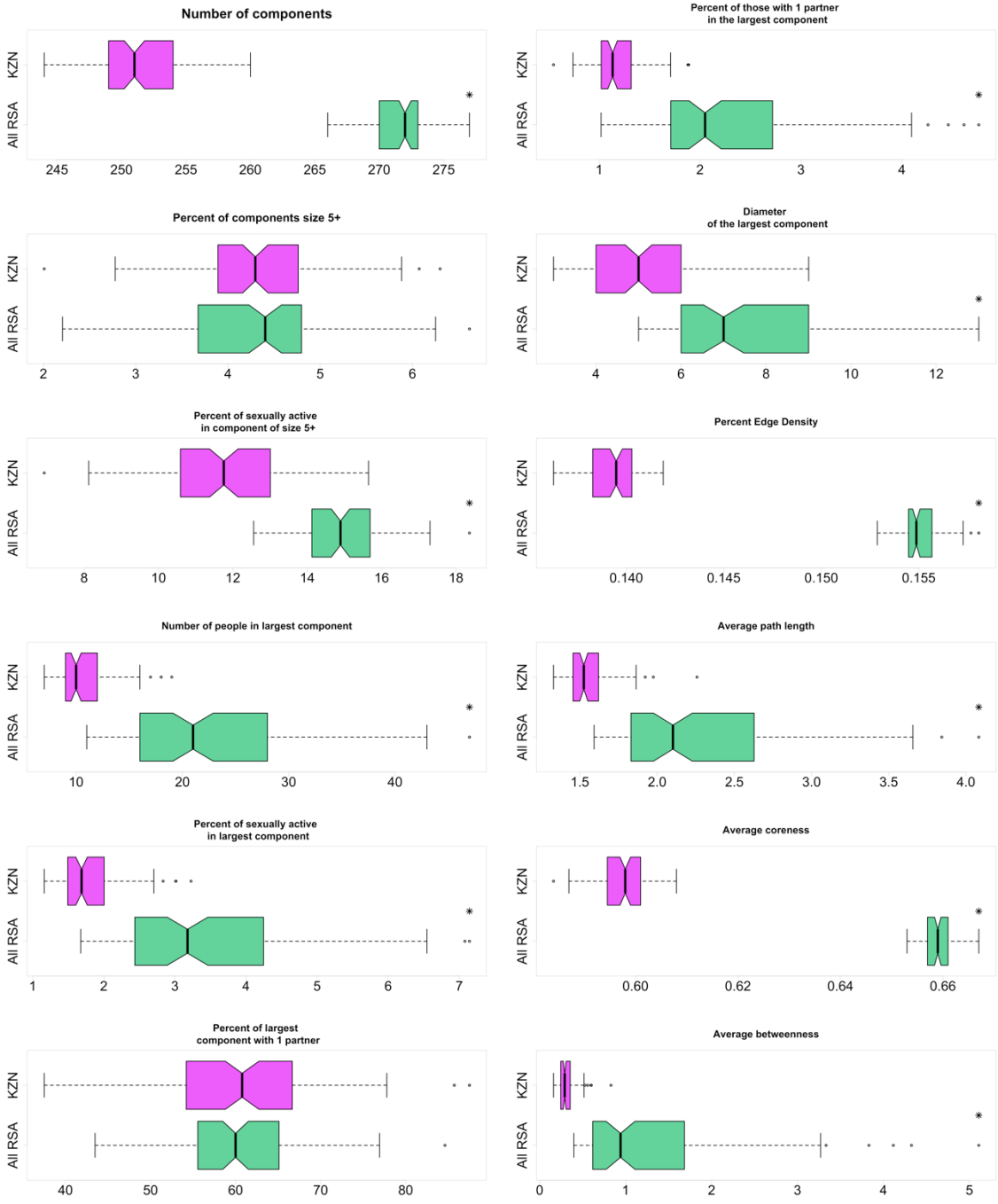
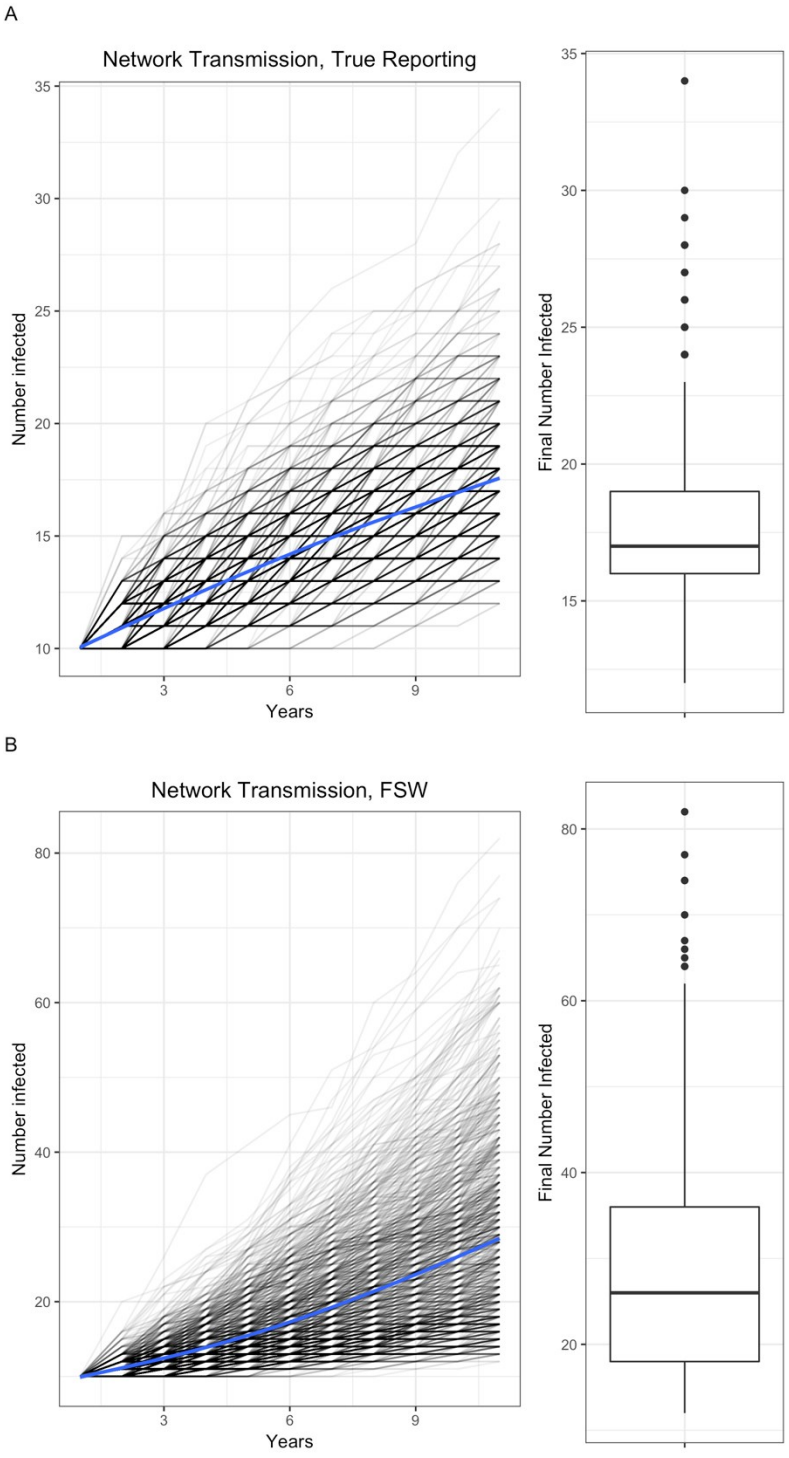


Figure S4.6: Full plot of trajectories for epidemic simulation (method A)



C

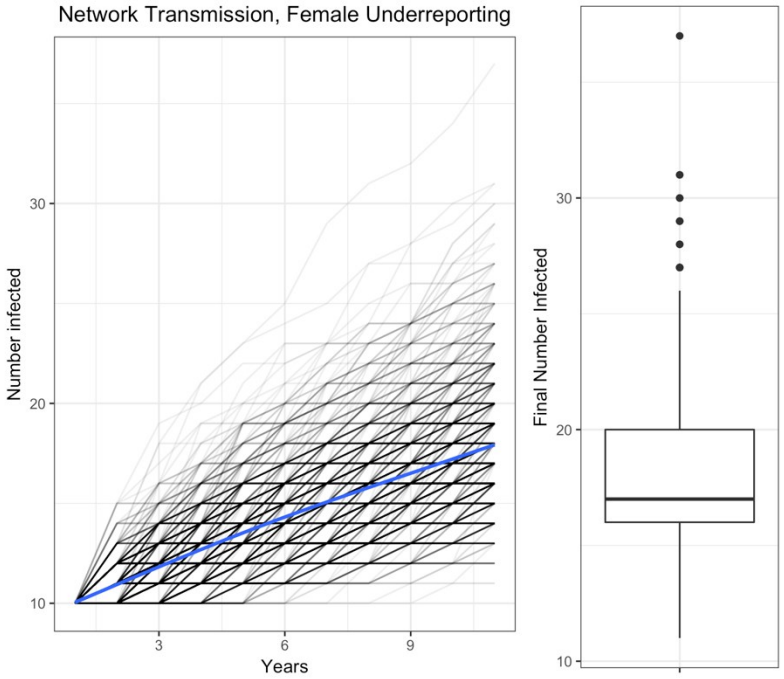
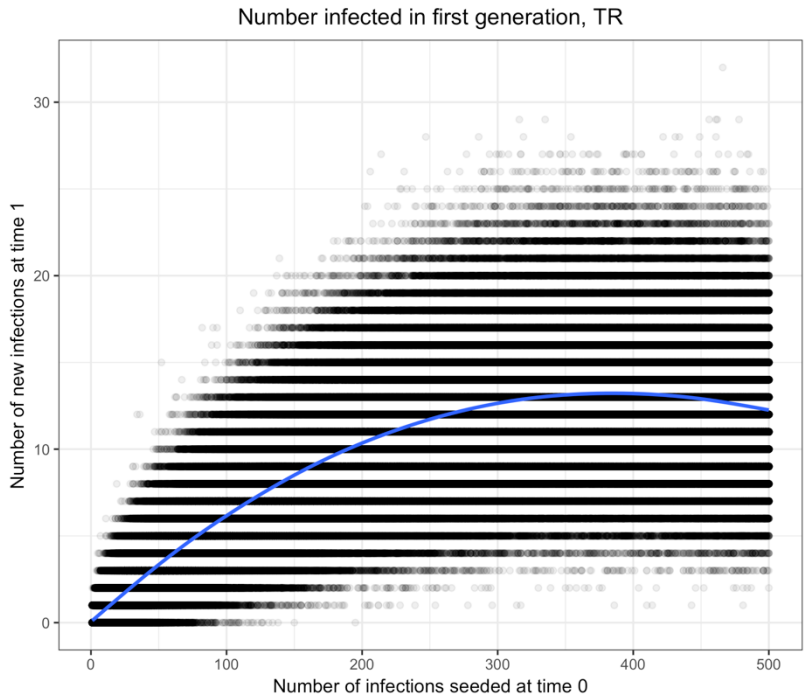
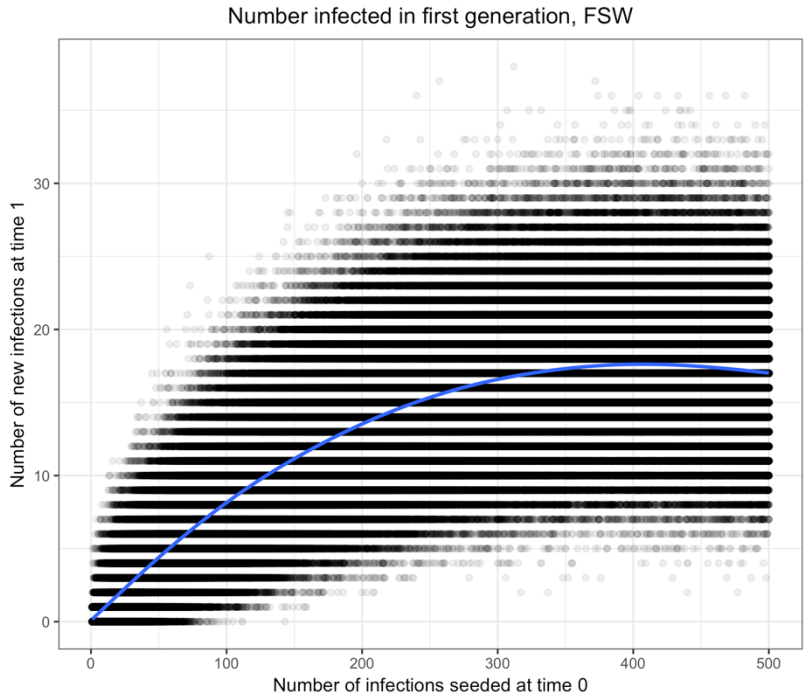

Network Transmission, Female Underreporting
Number infected
30
20
10
3
6
9
Years
Final Number Infected
30
20
10

Figure S4.7: Full plot of trajectories for epidemic simulation (method B)

A



B



C

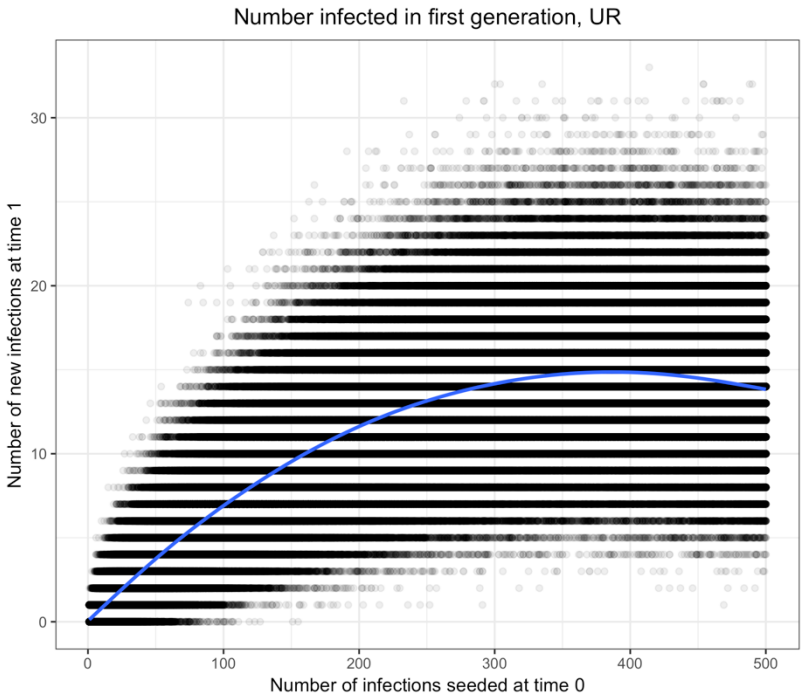
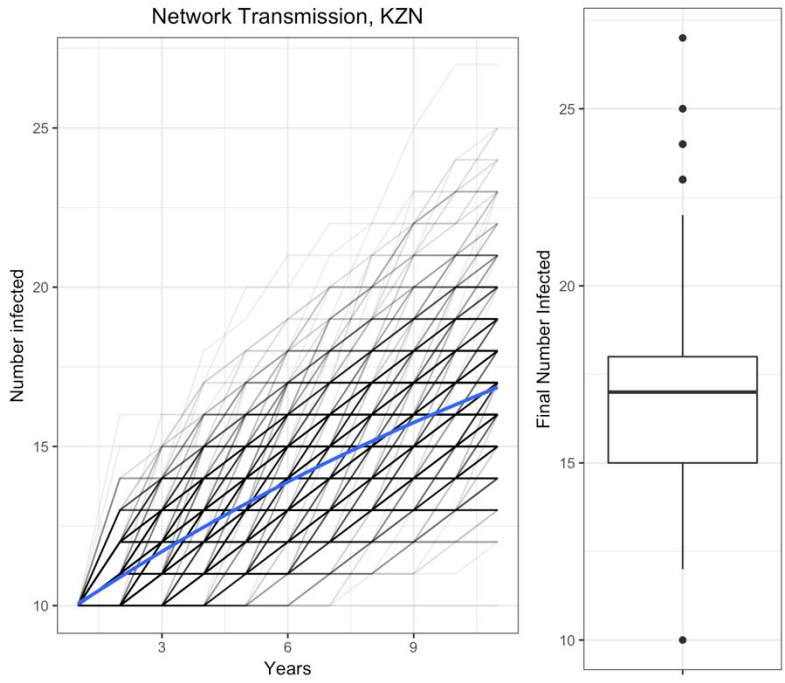
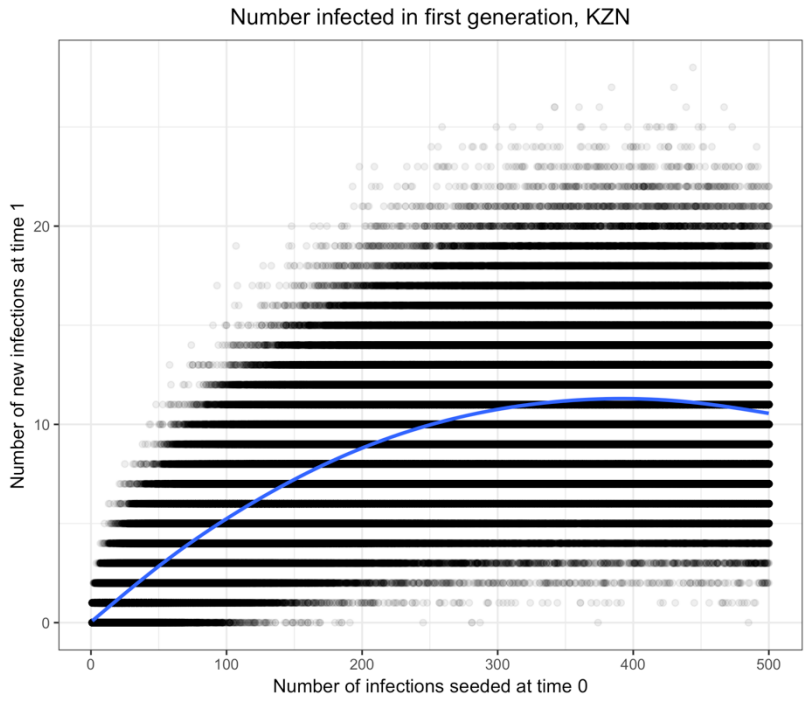

Number infected in first generation, UR
Number of new infections at time 1
0
10
20
30
0
100
200
300
400
500
Number of infections seeded at time 0

Figure S4.8: Full plot of trajectories for transmission potential in KwaZulu-Natal

A



B



4.7.6. Appendix References

1. Ramjee G, Williams B, Gouws E, Van Dyck E, De Deken B, Karim SA. The impact of incident and prevalent herpes simplex virus-2 infection on the incidence of HIV-1 infection among commercial sex workers in South Africa. Journal of acquired immune deficiency syndromes. 2005;39(3):333-9.

2. Merrigan MB, Tafuma TA, Okui LA, Lebelonyane R, Bolebantswe JM, Makhaola K, et al. HIV Prevalence and Risk Behaviors Among Female Sex Workers in Botswana: Results from the 2012 HIV/STI Bio-Behavioral Study. AIDS and behavior. 2015;19(5):899-908.

3. Chohan V, Baeten JM, Benki S, Graham SM, Lavreys L, Mandaliya K, et al. A prospective study of risk factors for herpes simplex virus type 2 acquisition among high-risk HIV-1 seronegative women in Kenya. Sexually transmitted infections. 2009;85(7):489-92.

4. Matovu JK, Ssebadduka BN. Sexual risk behaviours, condom use and sexually transmitted infection treatment-seeking behaviours among female sex workers and truck drivers in Uganda. International journal of STD & AIDS. 2012;23(4):267-73.

5. Konstant TL, Rangasami J, Stacey MJ, Stewart ML, Nogoduka C. Estimating the number of sex workers in South Africa: rapid population size estimation. AIDS and behavior. 2015;19 Suppl 1:S3-15.

6. South African National AIDS Council. Estimating the size of the sex worker population in South Africa, 2013 2013 Available: http://www.sweat.org.za/wp-content/uploads/2016/02/Sex-Workers-Size-Estimation-Study-2013.pdf.

7. Grabowski MK, Lessler J, Redd AD, Kagaayi J, Laeyendecker O, Ndyanabo A, et al. The role of viral introductions in sustaining community-based HIV epidemics in rural Uganda: evidence from spatial clustering, phylogenetics, and egocentric transmission models. PLoS medicine. 2014;11(3):e1001610.

8. Patel P, Borkowf CB, Brooks JT, Lasry A, Lansky A, Mermin J. Estimating per-act HIV transmission risk: a systematic review. Aids. 2014;28(10):1509-19.

9. 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.

10. Shisana O, Rehle T, Simbayi LC, Parker W, Zuma K, Bhana A, et al. South African National HIV Prevalence, HIV Incidence, Behaviours and Communication Survey. Cape Town: HSRC Press; 2005.

11. van Loggerenberg F, Mlisana K, Williamson C, Auld SC, Morris L, Gray CM, et al. Establishing a cohort at high risk of HIV infection in South Africa: challenges and experiences of the CAPRISA 002 acute infection study. PloS one. 2008;3(4):e1954.

5. Conclusions

In this dissertation, we examine sexual behavior, sexual mixing patterns and sexual networks derived from a nationally representative survey of South Africans conducted in 2012. Our results suggest that there is a tremendous amount of heterogeneity across the South African population in sexual behavior and sexual mixing and that this heterogeneity could have a great impact on HIV transmission. In South Africa, the country bearing the world's largest burden of HIV[1], our findings suggest there are several areas where an understanding of sexual behavior, sexual mixing and sexual networks could inform the implementation of HIV interventions.

In Chapter 2, we find that sexual behavior is more conservative among individuals living with HIV on ART, compared to among those living with HIV not on ART. This contributes to other evidence suggesting the importance of individuals living with HIV not on ART contributing disproportionately to HIV transmission[2]. Our findings suggest the importance for interventions to target individuals living with HIV but not on ART to engage them not only in HIV care and treatment but also in risk reduction interventions.

We show sexual assortativity along a number of different characteristics and characterize sexual partner degree distributions in Chapter 3. Among household partners, we see some evidence of assortativity by degree and ART status, as well as dramatic assortativity by race and age. Previous theoretical work has shown the importance of mixing patterns on epidemic trajectories, but has shown that the influence of assortativity depends on the stage of the epidemic and a variety of other characteristics[3, 4]. An

important next step to take using these results is to assess the impact that the documented levels of assortativity would have on an epidemic at South Africa's epidemic stage when we implement future HIV combination prevention interventions.

In Chapter 4, we present a novel nonparametric method for generating full network characteristics using limited egocentric and household partnership data. The methods allow for us to synthesize the space of sociometric networks that are consistent with these limited network data, without the extensive effort and costs inherent to a full sociometric study. We find that our method produces consistent networks that show a relatively narrow degree of variability on higher order network statistics, such as component size, average path length, or average betweenness. However, when we compare methods to balance male and female sexual partnership numbers, we find that the method used to balance these distributions matters a great deal. We find that if we assume all excess male partnerships are among sex workers, we find that sexual networks become much more highly connected and that the network can sustain a higher number of HIV cases, when compared to a network where we assume that all females have a consistent level of underreporting. Our results reinforce the importance of capturing high risk individuals, like sex workers, in serosurvey data in order to truly understand HIV transmission. It also suggests that the effectiveness of interventions to reduce HIV in populations will vary by our understanding of the spread of HIV, and that therefore this understanding should inform the selection of intervention approaches. In networks influenced dramatically by sex workers we will need to have a particularly targeted approach, but in the absence of a sex working or other much higher risk population our efforts will be more diffuse across the population.

This dissertation had a number of both strengths and weaknesses. Its largest weakness is the reliance on self-reported sexual behavior data for almost all of our findings. As self-reported data has been documented to have a range of biases[5-9], but biases that do not consistently trend in one direction or the other, our results are limited by the necessity to assume that self-reports are largely accurate. We do assess the impact of differential reporting by sex on our sexual networks in Chapter 4, which provides a small measure of assessment of the strength of our assumption. We additionally are limited by our exclusion of same-sex partnerships, excluded due to potential for bias. Same-sex partnerships (particularly men who have sex with men and transgender women) likely mix with and influence generalized heterosexual epidemics like that of South Africa[10]. Further, our study design likely does not capture migrant workers (while the single-sex hostels that migrant workers frequently live in were not explicitly excluded, the study was not designed to incorporate this population) or prisoners, who are explicitly excluded in the study design (SABSSM IV included non-institutionalized populations). Our study also has a number of strengths. First, we utilize a nationally representative sample of South Africans, which provides a wealth of data to characterize individuals across the country. Second, we utilize novel methods to generate sexual networks in Chapter 4 which expand upon previous modeling methods[11]. The nonparametric approach that we utilize gives nuanced understanding of sexual networks and can be extended to additional situations and fitting data. This method allowed for a principled assessment of balancing degree scenarios, an area that has largely been ignored by the body of network modeling research with a few notable exceptions[12].

Sexual networks can influence the potential effectiveness of combination HIV prevention interventions[13, 14]. In a time when research and programmatic dollars are ever decreasing, and the continuation of vital programs such as the President's Emergency Plan for AIDS Relief (PEPFAR) are in doubt[15], implementing programs in the way that will maximize our impact while minimizing costs is paramount. Using our improved understanding of the HIV epidemic based on sexual behavior and sexual networks to implement interventions that can have the greatest impact on curbing transmission is a key step at this time. Previous optimism regarding the potential for HIV treatment as prevention (TasP) to rapidly eliminate HIV transmission[16] has faced challenges in the context of difficulties linking individuals in South Africa to care[17, 18], a necessary second step for any treatment as prevention policy. Next steps include using our simulated sexual networks to assess network structure impacts on the effectiveness of HIV TasP, based on our understanding of differential uptake of TasP by the same characteristics that describe sexual mixing and networks. Using the findings from this dissertation and the body of work upon which it builds, we can design and implement interventions with the greatest impact to those most in need to combat South Africa's HIV epidemic.

5.1. References

1. UNAIDS. Global AIDS Update. 2016. Available from: http://www.unaids.org/sites/default/files/media_asset/global-AIDS-update-2016_en.pdf. Accessed 1 Feb 2017.

2. Skarbinski J, Rosenberg E, Paz-Bailey G, et al. Human immunodeficiency virus transmission at each step of the care continuum in the United States. JAMA Intern Med. 2015;175(4):588-96.

3. Gupta S, Anderson RM, May RM. Networks of sexual contacts: implications for the pattern of spread of HIV. AIDS. 1989;3(12):807-17.

4. Anderson RM, Gupta S, Ng W. The significance of sexual partner contact networks for the transmission dynamics of HIV. J Acquir Immune Defic Syndr. 1990;3(4):417-29.

5. 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:ii43-8.

6. Cleland J, Boerma JT, Carael M, Weir SS. Monitoring sexual behaviour in general populations: a synthesis of lessons of the past decade. Sex Transm Infect. 2004;80 Suppl 2:ii1-7.

7. Helleringer S, Kohler HP, Kalilani-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.

8. Mensch BS, Hewett PC, Erulkar AS. The reporting of sensitive behavior by adolescents: a methodological experiment in Kenya. Demography. 2003;40(2):247-68.

9. 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.

10. Middelkoop K, Rademeyer C, Brown BB, et al. Epidemiology of HIV-1 subtypes among men who have sex with men in Cape Town, South Africa. J Acquir Immune Defic Syndr. 2014;65(4):473-80.

11. Krivitsky PN, Handcock MS, Morris M. Adjusting for Network Size and Composition Effects in Exponential-Family Random Graph Models. Stat Methodol. 2011;8(4):319-39.

12. Merli MG, Moody J, Mendelsohn J, Gauthier R. Sexual Mixing in Shanghai: Are Heterosexual Contact Patterns Compatible With an HIV/AIDS Epidemic? Demography. 2015;52(3):919-42.

13. Delva W, Helleringer S. Beyond Risk Compensation: Clusters of Antiretroviral Treatment (ART) Users in Sexual Networks Can Modify the Impact of ART on HIV Incidence. PLoS One. 2016;11(9):e0163159.

14. Jenness SM, Goodreau SM, Morris M, Cassels S. Effectiveness of combination packages for HIV-1 prevention in sub-Saharan Africa depends on partnership network structure: a mathematical modelling study. Sex Transm Infect. 2016;92(8):619-24.

15. The Lancet H. Will President Trump protect his party's PEPFAR legacy? Lancet HIV. 2017;4(1):e1.

16. Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet. 2009;373(9657):48-57.

17. Iwuji C, Orne-Gliemann J, Balestre E, et al. The impact of universal test and treat on HIV incidence in a rural South African population: ANRS 12249 TasP trial, 2012-2016. International AIDS Conference; Durban, South Africa2016.

18. Haber N, Tanser F, Bor J, et al. From HIV infection to therapeutic response: a population-based longitudinal HIV cascade-of-care study in KwaZulu-Natal, South Africa. Lancet HIV. 2017.

June 6, 2017

6. CURRICULUM VITAE

KATHRYN A. RISHER, M.H.S.

PERSONAL DATA
Home Address
134 F St SE, Apt 3
Washington, DC 20003
Work Address
Department of Epidemiology, Rm W6604
Johns Hopkins Bloomberg School of Public Health
615 N. Wolfe Street, Baltimore, MD 21205
Email: krisher1@jhu.edu
Cell: (708) 334-9566

EDUCATION

PhD/Current (Exp. May 2017)	Johns Hopkins Bloomberg School of Public Health, Baltimore, MD (Infectious Disease Epidemiology). Thesis Title: "*Sexual behavior and sexual networks in South Africa: Implications for HIV transmission.*" Advisors: Shruti Mehta, Justin Lessler
MHS/Current (Exp. May 2017)	Johns Hopkins Bloomberg School of Public Health, Baltimore, MD (Biostatistics). Advisor: Elizabeth Ogburn
MHS/2013	Johns Hopkins Bloomberg School of Public Health, Baltimore MD (Infectious Disease Epidemiology). Thesis Title: "*Determinants of perceived and enacted stigma among men who have sex with men in Swaziland.*" Advisor: Stefan Baral
BA/2010	Dartmouth College, Hanover, NH (Biology, International Studies minor).

OTHER TRAINING

2015	Summer Institute in Statistics and Modeling of Infectious Diseases, University of Washington (2 weeks)

2014	Network Modeling for Epidemics, University of Washington (1 week)
2014	Meaningful Modeling of Epidemiological Data, International Clinics on Infectious Disease Dynamics and Data (2 weeks)

PROFESSIONAL EXPERIENCE

2013-2016	Research Assistant (3 mos full time, part time to 2016), Human Sciences Reserch Council, Cape Town, South Africa. *Supervisors:* Olive Shisana, Thomas Rehle, Leickness Simbayi
2011-2017	Research Assistant, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD. *Supervisors*: David Dowdy & Maunank Shah (2013-2016), Cecilia Tomori (2014), Wendy Davis & David Celentano (2012-2013), Stefan Baral (2011-2012)
2010-2011	Post-Baccalaureate Client Services Intern, UCSF AIDS Health Project (Now UCSF Alliance Health Project), San Francisco, CA. *Supervisor*: Emily Leavitt
2009	Health Promotion Intern (3 mos), Erie Family Health Center, Chicago, IL. *Supervisor*: Lauren Ranalli
2008	International Health Volunteer (3 mos), Global Service Corps, Arusha, Tanzania. *Supervisor*: Jenaya Rockman
2006-2007	Trade Support and Research Intern (6 mos – summers), Spot Trading, LLC, Chicago, IL. *Supervisor*: Bob Gabriel

PROFESSIONAL ACTIVITIES

Society Membership

2012-present	International AIDS Society
2014-present	Society for Epidemiologic Research

EDITORIAL ACTIVITIES

Peer Review Activities. Journal Reviewer

Journal of AIDS (JAIDS)
Open Forum Infectious Diseases (OFID)
PLoS One

HONORS AND AWARDS

2015	Louis I. and Thomas D. Dublin Award for the Advancement of Epidemiology and Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
2015-2016	Epidemiology Department 100% Scholarship and Stipend, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
2014	Charlotte Silverman Award, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD
2013-2015	Johns Hopkins HIV Epidemiology and Prevention Sciences Training Program (T32 AI102623), Baltimore, MD
2013	Delta Omega Public Health Honor Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

PUBLICATIONS

Peer Reviewed Journal Articles (denotes co-first authorship)*

1. **Risher K,** Adams D, Sithole B, Ketende S, Kennedy C, Mnisi Z, Mabusa X, Baral SD. Sexual stigma and discrimination as barriers to seeking appropriate healthcare among men who have sex with men in Swaziland. *J Int AIDS Soc.* 2013; 16(3 Suppl 2):18715.
2. Tomori C, **Risher K**, Limaye RJ, Van Lith LM, Gibbs S, Smelyanskaya M, Celentano DD. A role for health communication in the continuum of HIV care, treatment, and prevention. *J Acquir Immune Defic Syndr*. 2014; 66(Suppl 3): S306-10.
3. Shisana O, Zungu N, Evans M, **Risher K**, Rehle T, Celentano D. The case for expanding the definition of 'key populations' to include high-risk groups in the general population to improve targeted HIV prevention efforts. *S Afr Med J*. 2015; 105(8):669.
4. **Risher K**, Mayer KH, Beyrer C. HIV Treatment cascade in MSM, people who inject drugs, and sex workers. *Curr Opin HIV AIDS*. 2015; 10(6):420-9.
5. Shisana O, **Risher K**, Celentano DD, Zungu N, Rehle T, Ngcaweni B, Evans MG. Does marital status matter in an HIV hyperendemic country? Findings from the 2012 South African National HIV Prevalence, Incidence and Behaviour Survey. *AIDS Care*. 2016; 28(2):234-41.
6. Shah M, **Risher K**, Berry SA, Dowdy D. The epidemiologic and economic impact of improving HIV testing, linkage, and retention in care in the United States. *Clin Infect Dis*. 2016; 62(2):220-9.
7. Shah M, Perry A, **Risher K**, Kapoor S, Grey J, Sharma A, Rosenberg ES, Del Rio C, Sullivan P, Dowdy DW. Quantifying the impact of the National HIV/AIDS Strategy targets for improved HIV care engagement in the US. *Lancet HIV*. 2016; 3(3):e140-6.
8. **Risher K**, Rehle T, Simbayi L, Shisana O, Celentano DD. Antiretroviral treatment and sexual risk behavior in South Africa. *AIDS Behav*. 2016; 20(4):710-6.
9. Means AR*, **Risher K***, Ujeneza EL*, Maposa I, Nondi J, Bellan SE. Impact of age and sex on CD4+ cell count trajectories following treatment initiation: an analysis of the Tanzanian HIV treatment database. *PLoS One*. 2016; 11(10):e0164148.

10. Evans MGB, **Risher K**, Zungu M, Shisana O, Moyo S, Celentano DD, Maughan-Brown B, Rehle TM. Age-disparate sex and HIV risk for young women from 2002 to 2012 in South Africa. *J Int AIDS Soc*. 2016; 19(1):21310.
11. **Risher K**, Kapoor S, Daramola AM, Paz-Bailey G, Skarbiski J, Doyle K, Shearer K, Dowdy D, Rosenberg E, Sullivan P, Shah M. Challenges in the evaluation of interventions to improve engagement along the HIV care continuum in the United States: a systematic review. *AIDS Behav*. 2017.
12. Chersich M, Wabiri J, **Risher K**, Shisana O, Celentano D, Rehle T, Evans M, Rees H. Contraception coverage and methods used among women in South Africa: findings of a national household survey. *S Afr Med J*. In Press.

CURRICULUM VITAE

KATHRYN A. RISHER, M.H.S.

PART II

TEACHING

Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

2012 Epidemiology and Natural History of Human Viral Infections (340.654), Teaching Assistant, 25 graduate students. Supervisor: Homayoon Farzadegan

2012 Epidemiology and Public Health Impact of HIV/AIDS (340.646), Teaching Assistant, 30 graduate students. Supervisor: Homayoon Farzadegan

2013 Advanced Topics on Control and Prevention of HIV/AIDS (340.744), Teaching Assistant, 18 graduate students. Supervsior: Homayoon Farzadegan

2013 Concepts and Methods in Infectious Disease Epidemiology (340.609), Teaching Assistant, 25 graduate students. Supervisors: Justin Lessler, Derek Cummings, Bill Moss

2013 Principles of Epidemiology (340.601), Teaching Assistant, 150 graduate students. Supervisors: David Celentano, Bill Moss

2013 Epidemiologic Methods II (340.752), Teaching Assistant, 125 graduate students. Supervisors: Gypsyamber D'Souza, Stephan Ehrhardt

2015 Observational Epidemiology (340.608), Teaching Assistant, 60 graduate students. Supervisors: Stephen Gange, Elizabeth Golub

RESEARCH GRANT PARTICIPATION

9/1/16-2017 Impact of Heterogeneity in Structural and Network Level Factors on HIV Treatment as Prevention in South Africa
F31AI120780
$43,576
NIAID
Risher (PI): Predoctoral Ruth L. Kirschstein National Research Service Award to support doctoral training and dissertation work conducting analysis and network modeling with the Human Sciences Research Council's 2012 household survey

ACADEMIC SERVICE

Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

2015-2016	Departmental Centennial Committee, Student Representative
2014-	Surveillance and Outbreak Response Team, Member
2013-2014	Epidemiology Student Organization, Co-President
2013-2014	Infectious Disease Epidemiology Research in Progress, Student Coordinator
2012-2013	Infectious Disease Epidemiology Journal Club, Student Coordinator

PRESENTATIONS

Presentations at Scientific Meetings (Oral Presentations)

2015 **Risher K**, Shah M, Paz-Bailey G, Wejnert C, Shouse RL, Dowdy D, Sullivan P, Rosenberg E, Skarbinski J. Comparison of interventions across the HIV care continuum. National HIV Prevention Conference, Atlanta, USA.

Presentations at Scientific Meetings (Posters, presented on site)

2014 Rehle T, Simbayi L, Shisana O, Celentano D, **Risher K**. Sexual behavior among people living with HIV/AIDS who are on antiretroviral therapy in South Africa: results from a nationally representative household survey in 2012. XX International AIDS Conference, Melbourne, Australia.

2014 Celentano D, **Risher K**, Shisana O, Rehle T, Simbayi L. Prevalence, correlates and ecological associations of concurrent sexual partnerships in a nationally representative cross-section of South Africans in 2012. XX International AIDS Conference, Melbourne, Australia.

2015 **Risher K**, Lessler J, Shisana O, Rehle T, Simbayi L, Evans M, Celentano DD. Sexual mixing patterns by age in a nationally representative sample of South Africans: implications for HIV transmission. Epidemics 5: Fifth International Conference on Infectious Disease Dynamics, Clearwater Beach, USA.