Stochastic Network Models for HIV-1 Transmission Dynamics

Samuel M. Jenness

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:

Susan Cassels, Chair

Steven Goodreau

Martina Morris

Program Authorized to Offer Degree:

Public Health – Epidemiology

University of Washington

Abstract

Stochastic Network Models for HIV-1 Transmission Dynamics

Samuel M. Jenness

Chair of the Supervisory Committee:

Professor Susan Cassels

Department of Epidemiology

The studies in this dissertation investigate how the structure of dynamically evolving sexual networks shape the HIV-1 epidemic among heterosexuals in Sub-Saharan Africa (SSA). The aims of this project were to: 1) develop a comprehensive set of demographic, behavioral, and biological parameters characterizing the target population of heterosexuals in SSA for use in network-based mathematical models of HIV transmission dynamics; 2) mathematically model the synergistic effects of network structure and male circumcision on HIV transmission in SSA; and 3) use mathematical modeling to estimate the total, direct, and indirect effects of pre-exposure prophylaxis (PreP) on HIV incidence within a simulated randomized control trial (RCT) environment with counterfactual scenarios for threshold levels of network structures that can bias the estimation of treatment efficacy.

The heterosexual spread of HIV-1 infection in SSA depends on the unique configurations of how sexually active persons form and break sexual partnerships over time. For disease transmission, individuals are linked in dyads through partnerships, dyads are connected to other dyads when persons have multiple ongoing partnerships, and this forms the larger sexual network within the population. Effective HIV prevention tools, like male circumcision and PreP, operate within this network context. The studies here ask how these interventions, targeted at individuals, function given dynamic networks. Stochastic network models were developed to test key hypotheses for this interaction, aimed both at the

population level and within RCT settings. Parameters for these models were primarily drawn from an original retrospective panel study we conducted in Accra, Ghana specifically for mathematical modeling.

The findings from these studies address important empirical questions on the relationship between biomedical prevention tools and socio-behavioral risk, and also provide insight into the design and targeting of single-element and combination prevention packages for HIV in high-incidence settings. The broader mathematical modeling methods in this project have many potential applications for future HIV prevention research.

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

ACKNOWLEDGEMENTS

I would sincerely like to thank my dissertation committee for all their guidance and feedback on this research project: Susan Cassels, Steven Goodreau, Martina Morris, Jared Baeten, and Thomas Richardson. Special thanks go to Susan for her close mentoring over the past 3 years. My lab group, the UW Network Modeling Group, has served as both a nurturing incubator and a critical sounding board during my time at UW. Thank you to all the group members who have reviewed prior iterations of this research. I would like to especially thank Aditya Khanna for his feedback on my ideas over the years. The Migration & HIV in Ghana study provided a unique, valuable research experience, both personally and professionally, and it generated the empirical data for this project. Many thanks go to our scientific collaborators in Ghana: Adriana Biney, William Ampofo, and Francis Dodoo. The network models in these studies depend heavily on the statistical software authored by the Statnet Development Team. Thank you to Steve and Martina for teaching me networks, and to Skye Bender-deMoll for his coding guidance and software support. I was able to complete this dissertation project as quickly as I did through the support of a pre-doctoral fellowship at the UW Centers for the Studies in Demography and Ecology (CSDE). I would like to thank Sara Curran and Matt Weatherford at CSDE in particular for their support over the past 2 years. This work is funded by a number of research, training, and infrastructure grants: R00 HD057533 (NICHD; PI Susan Cassels); P30 AI027757 (UW CFAR; PI Susan Cassels); R01 HD68395 (NICHD; PI Martina Morris); T32 HD007543 (NICHD; PI Mark Ellis); R24 HD042828 (NICHD; PI Mark Ellis). Finally, I would like to express my most grateful thanks to friends and family, but Leah and Miles most of all.

INTRODUCTION

The heterosexual spread of HIV-1 infection in Sub-Saharan Africa (SSA) depends on the unique patterns of how sexually active persons form and break sexual partnerships over time, along with the biological determinants of HIV transmission and acquisition within these partnerships. Sexual partnership networks comprise the complete set of partnership dyads within a population. The connectivity of the network, which is partially a function of partnership degree (how many partnerships persons have at any one time), greatly impacts the speed and magnitude with which HIV circulates within a population. Factors influencing the transmission rate within partnerships include the HIV viral load of the infected partner, as well as the circumcision status, co-morbid sexually transmitted infections, age, and sex of the susceptible partner.

This dissertation investigates three related questions at the intersection between biomedical prevention tools and structure of sexual networks. The studies in this dissertation investigate both the factors generating the network structure, and also what leads to an HIV transmission within a partnership given that structure.

Question 1

How Does Degree Impact the Per-Partnership HIV Transmission Rate?

The first aim of this dissertation was to develop a comprehensive set of demographic, behavioral, and biological parameters characterizing our target population for use in network-based mathematical models of HIV transmission dynamics. Our target population that motivates this research generally is heterosexuals within SSA. However, we started with a high-risk population in Ghana that served as the basis for our empirical study, the Migration & HIV in Ghana (MHG) study. MHG was a cross-sectional study of sexually active adults in Agbogbloshie, Ghana conducted in 2012. Agbogbloshie is an urban resource-poor area in the capital city of Accra, selected based on its hypothesized high-risk profile and lack of prior research. As outlined in Chapter 1, we conducted a population-based study that collected data on the demographics and sexual network structure, along with HIV serostatus, of this population.

The comprehensive parameter set for the mathematical models is outlined in detail in the supplemental appendix of the first mathematical modeling paper, Chapter 2. Because this descriptive aggregation of data did not address a scientific question per se, Chapter 1 pursues a specific question related to mathematical model inputs using more traditional statistical data analysis of our empirical data from Ghana.

The question of how degree impacts the per-partnership HIV transmission rate concerns whether people reduce their coital frequency within a partnership when they add an additional concurrent partner. Under full coital dilution, persons keep their overall sexual act rate across partners constant, regardless of the number of partners they have. Under no dilution, coital frequency is a constant multiple of the number of partners at any one time. Models have suggested that coital dilution could effectively

eliminate the disease accelerating effects of concurrency, but empirically, the data are mixed on whether dilution actually occurs in SSA. If dilution *does* occur, then the per-partnership HIV transmission rate would by modified by degree, with the rate needing to be adjusted downward to account for the fewer acts per partnership. Chapter 1 investigates this association between partnership degree and both total and unprotected sexual acts, using time-varying exposure and outcome measures collected in retrospective panel data from the MHG study. The goal is to quantify dilution in western Africa to add to prior dilution data from eastern and southern Africa, and to support the mathematical modeling with dilution parameters as necessary.

Question 2

What Conditions of Network Structure Maximize the Effectiveness of Male Circumcision?

The second aim of this project was to mathematically model the synergistic effects of network structure and male circumcision on HIV transmission in SSA. Chapter 2 investigates how components of the dynamic network structure influence the effectiveness of male circumcision as a tool to prevent HIV-1 infection.

Three randomized controlled trials (RCTs) of male circumcision independently observed an HIV prevention benefit of over 50%, but a subsequent trial found that circumcised HIV-infected men did not find a risk reduction to uninfected female partners. Despite this, mathematical models have predicted long-term benefits of male circumcision on HIV incidence among women, driven by the reduction in incidence in men. Ecologically, differences in circumcision prevalence have been suggested to explain the large regional disparity in HIV-1 burden across SSA: male circumcision is nearly universal in many western African counties where disease prevalence is much lower than southern African counties, where circumcision is less common. Prior studies suggest that a large-scale reduction in HIV incidence would require circumcising a large fraction of the population. Circumcision campaigns to date have been successful, but have not reached their targets. Prevention packages that combine circumcision with behavioral intervention elements may yield a large impact on incidence while requiring a smaller scale-up.

In this study, we investigate how changes to the network connectivity synergistically acted with changes to circumcision to reduce HIV incidence among heterosexuals in SSA. Two related questions are answered. First, does the varying prevalence of circumcision partially explain the regional disparities in the HIV-1 epidemic in SSA? Second, what are the synergistic effects of changes to the network structure and circumcision coverage? To investigate this, we use stochastic network-based mathematical models for HIV-1 transmission dynamics. Compared to prior mathematical models examining circumcision and behavioral synergy, our network models do not require extreme assumptions about baseline behavior or changes to it.

Question 3

Can Network Structure Bias Estimates of Pre-Exposure Prophylaxis Efficacy within Clinical Trials?

The third aim is to use mathematical modeling to estimate the total, direct, and indirect effects of pre-exposure prophylaxis (PrEP) on HIV incidence within a randomized controlled trial (RCT) setting. RCTs of PrEP may be subject to interference, which occurs when the probability of infection for a study participant depends on the treatment status not just of that individual but also of others in the study and population from which they were recruited. Interference generates indirect effects for an intervention by conferring prevention benefits to those in the control arm. The effects of interference will depend on how the infectious disease under investigation circulates within the community, which itself depends on the contact patterns for disease transmission. The interference literature to date has not investigated contact structures with dynamic and complex mixing patterns, such as sexual partnerships. Ours will be one of the first to quantify interference within studies of HIV and sexually transmitted infection prevention agents using empirical data on behavior.

Our motivating case for this study is oral PrEP among heterosexuals in SSA. While oral PrEP has generally been shown to be an effective agent for the prevention of HIV-1 infection, there have been discrepancies in RCT results. The Fem-PrEP and VOICE trials in particular have observed no disease prevention benefit. Adherence to the study drug, measured by pill count and serum drug concentration, was low in the null trials, suggesting that the intent-to-treat estimator provides little information about the true efficacy of PrEP. However, there are also differences in the trial designs that may help explain the variability. Null trials like Fem-PrEP were conducted among young women with high rates of partner turnover and concurrency, whereas more effective trials like Partners PrEP were among stable serodiscordant couples in long-term partnerships.

The study in Chapter 3 uses mathematical modeling to investigate how interference can bias the observable results of a study environment like Fem-PrEP. The hypothesis is that the structural elements of the sexual network in a high-incidence setting like this will recirculate disease rapidly enough to confer non-trivial indirect effects that reduce the observable differences in incidence between treatment and control arms, diluting the hazard ratio. Our mathematical modeling framework allows for testing different components of the study design and setting that contribute to this bias. The broader goal of this research is to provide a more general exposition of a model-based approach to the partitioning of total effects into direct and indirect effects for bias estimation in future prevention research.

Chapter 1: Minimal Coital Dilution in Accra, Ghana

Published citation:

Jenness SM, Biney AA, Ampofo WK, Nii-Amoo Dodoo F, Cassels S. Minimal Coital Dilution in Accra, Ghana. *Journal of acquired immune deficiency syndromes.* Epub ahead of print: Jan 23 2015.

Title

Minimal Coital Dilution in Accra, Ghana

Authors

Samuel M. Jenness, MPH[1]
Adriana A. E. Biney[2]
William Kwabena Ampofo[3]
F. Nii-Amoo Dodoo[2, 4]
Susan Cassels, PhD[1, 5]

[1] Department of Epidemiology, University of Washington, Seattle, WA
[2] Regional Institute for Population Studies, University of Ghana, Legon
[3] Noguchi Memorial Institute for Medical Research, University of Ghana, Legon
[4] Pennsylvania State University, University Park, PA
[5] Department of Geography, University of California, Santa Barbara, CA

Correspondence

Samuel M. Jenness, Department of Epidemiology, University of Washington. 1959 Pacific Street NE, Box 357236, Seattle, WA 98195. *Email:* sjenness@uw.edu.

Sources of Support

This work was supported in part by the NICHD (R00 HD057533) and the UW Center for AIDS Research SPRC (P30 AI027757). Computing support was provided by a NICHD research infrastructure grant (5R24HD042828) to the UW Center for Studies in Demography & Ecology.

Running Head

Minimal Coital Dilution in West Africa

Word Counts

Manuscript: 3485 words
Abstract: 222 words

ABSTRACT

Background Coital dilution, the reduction in the coital frequency per partner when an additional ongoing partner is added, may reduce the transmission potential of partnership concurrency for HIV and other sexually transmitted infections. Empirical estimates of dilution, especially dilution of sexual acts unprotected by condoms, are needed to inform prevention research.

Methods Sexually active adults in Accra, Ghana were recruited in a multi-stage household probability sample. Degree (number of ongoing partners), total acts, and unprotected acts were measured retrospectively for each month in the past year through an event-history calendar. Random effects negative binomial models estimated the association between degree and coital frequency.

Results Compared to person-months with a single partner (monogamy), 2.06 times as many total acts and 1.94 times as many unprotected acts occurred in months with 2 partners. In months with 3 partners, 2.90 times as many total acts and 2.39 times as many unprotected acts occurred compared to monogamous months. Total acts but not unprotected acts also declined with partnership duration.

Conclusions No dilution was observed for total acts with up to three concurrent partners, but a small amount of dilution was observed for unprotected acts for months with multiple concurrencies. This suggests moderate selective condom use in months with multiple concurrencies. The implications of the observed dilution for future HIV transmission must be investigated with mathematical models.

KEY WORDS

Coital dilution; coital frequency; concurrency; HIV/AIDS; Ghana; West Africa

INTRODUCTION

Coital dilution is the reduction in the number of sexual acts per partner when an additional, concurrent partner is added [1]. Concurrency is defined as having two or more sexual partnerships overlapping in time [2]. Figure 1 shows one extreme, in which total (100%) coital dilution would occur if a person who has a primary partner and a coital frequency of 4 acts per month acquires a secondary partner, but splits the 4 acts between the two partners (equally, or in any combination preserving the total across partners). At the other extreme, 0% dilution would occur if this person now has twice as many acts, so the number per partner remains constant. While this metric does not allow for increasing activity per partner, this is also certainly possible (e.g., a long-term partnership supplemented by a new partnership with sufficiently high coital frequency).

Dilution may mitigate the effects of concurrency for the transmission of HIV and other sexually transmitted infections. Concurrency has been shown to increase disease incidence by allowing for the virus to transmit backwards from those partners established later to those earlier, and by decreasing the length of the generation interval [2]. But dilution decreases the per-partnership transmission rate, since this is a function of the number of sexual acts in each partnership per unit time. A common assumption in mathematical models is to preserve the number of acts per partnership under concurrency [3], but that implies the number of potential transmission events increases linearly with degree (the number of ongoing partners at any time). One model testing this assumption suggested that relatively modest levels of dilution (25% or higher) could erase the effects of concurrency [1], though the assumptions of this model also led to decreasing rates of activity over time.

The critical empirical question, therefore, is whether dilution occurs to the level required to offset the effects of concurrency. Despite scientific interest in concurrency, few empirical studies have addressed dilution. Modest dilution was found within the context of polygamous marriages, but the prevalence of these partnerships and their epidemic potential is minimal in many African populations [4]. Another study investigated concurrency in Kenya and Malawi using various analytic approaches, finding evidence for dilution in some cases but not others [5]. A review of three studies in Uganda, Thailand, and the United States found consistently lower rates of coital frequency with secondary compared to primary partners, with that difference the smallest in the country with the highest prevalence (Uganda) [6].

One aspect of dilution requiring attention is condom use. Given the effectiveness of condoms for HIV/STI prevention [7], the most epidemiologically relevant dilution involves unprotected acts. Unprotected dilution occurs even if the total acts (protected and unprotected) increase with degree as long as the number of unprotected acts across partners remains constant. If 10 unprotected acts with a primary partner are supplemented with 10 protected acts with a secondary partner, no dilution has occurred with respect to total acts, but 100% dilution

has occurred for unprotected acts. In South Africa, condom use did not vary by concurrency [8], but more precise quantification of the association between degree and unprotected acts is still needed. The impact of unprotected dilution on disease transmission is complex, since condom use may vary by partner type and partnership duration [9].

In this study, we investigated the association between partnership degree and both total and unprotected sexual acts, using time-varying exposure and outcome measures collected in retrospective panel data. As our target population for this study was sexually active adults in a dynamic urban slum area in Accra, Ghana [10], this is the first study of dilution in West Africa, which has experienced less HIV-1 burden compared to other African regions [11]. Our goal is to quantify dilution there in order to support future mathematical modeling of HIV transmission dynamics to investigate these regional disease disparities.

METHODS

Procedures. This analysis uses data from the Migration & HIV in Ghana (MHG) study, a cross-sectional study of sexually active adults in Agbogbloshie, Ghana in 2012. Agbogbloshie is an urban slum area in the capital city of Accra, selected for this study based on its hypothesized high-risk profile and lack of prior epidemiological research. The methods have been described in detail [12]. Briefly, MHG used a two-stage cluster randomized sampling scheme to obtain a probability sample of the population. Starting with an area census, we first randomly selected households with a probability proportional to household size, and then randomly selected one adult household member. Given differences in household size, a weighting scheme was employed to account for differential inclusion probabilities. Eligibility criteria to participate were current residence in the selected household, age 18 to 49 years, and lifetime history of consensual sexual intercourse. The Institutional Review Boards of the University of Washington and University of Ghana approved all procedures.

Measures. Trained field staff administered a standardized survey and drew serum via finger stick for a diagnostic HIV-1/2 test. The survey focused on demographics, migration and travel, and sexual behavior. For sexual behavior, summary data were collected on the number of lifetime sexual partners, past-year partners, and past-year partners with whom condoms were not always used. We used an event-history calendar to collect detailed partnership data for partners in the past year (up to three), with responses for each month during that period [13]. For each partner, data included the duration of the partnership and monthly information on the number of total and unprotected sexual acts. The degree for each person-month is defined as the number of overlapping partnerships in that month (up to three). To be categorized as a degree-2 month, for example, acts with the second partner must have occurred between those with the first partner. Ambiguous months, such as those in which one partnership ended and another started, were conservatively categorized as with a single partner (16 out of 4465 person-months).

For the HIV testing, dried blood spots were collected in the field on standard filter paper and maintained in refrigerators before delivery to the Department of Virology, Noguchi Memorial Institute for Medical Research, University of Ghana for processing. Serum was tested on the INNO-LIA^{TM} HIV-1/2 test platform (Innogenetics, Belgium), shown to have good sensitivity and specificity for diagnosis and HIV type differentiation. All subjects who tested for HIV were asked to return in one week to receive their test results, which were provided by a trained nurse counselor. HIV-infected subjects were referred to medical care.

Statistical Analysis. Descriptive statistics were estimated for the full sample of sexually active adults in the past year. Because past-year sexual activity was not a study eligibility criterion, the analytic sample therefore excludes subjects inactive during the entire year. To account for the complex survey design, data were weighted by the inverse of the selection probability given the cluster sampling scheme, necessary because a discrepancy in household sizes from the initial census was observed [12]. Robust cluster-based standard errors were used for variance estimation [14]. Statistics for degree and monthly acts were calculated on person-months, while statistics that are temporally invariant are calculated on persons.

For the primary analysis, we investigated coital dilution by regressing sexual acts on degree with two random-effects negative binomial models. The outcome in the first model was the number of total acts summed across all partners (up to 3) in each person-month. In the second model, the outcome was unprotected acts across partners. The main predictor in both models was degree in each person-month, parameterized as an indicator variable with the levels measured in the survey (1, 2, or 3 partners). Two additional variables were selected a priori as precision variables: sex of the subject and the duration of the longest-running partnership in that month. The duration variable was included to allow the predicted acts to vary for long-term partnerships. Concurrency in sub-Saharan Africa often occurs within the context of these partnerships [6], and we expected fewer acts in those established partnerships.

The random effects parameterization nested person-months within persons. Person-months with a degree of zero were dropped from the analysis as coital frequency must always be zero in these months. The negative binomial response distribution allowed for overdispersion in the outcomes. Exponentiated coefficients from these models were interpreted as incidence rate ratios for degree conditional on sex and duration. If 100% coital dilution holds, then these ratios for both degree comparisons (2 vs. 1 and 3 vs. 1) should be one. If 0% dilution holds, then the ratios should be a linear function of degree: degree-2 person-months should have twice as many acts as degree-1 months, and degree-3 should be three times as high. Finally, we estimated the predicted number of acts in two ways: first by degree and sex averaged over duration, and second by degree and duration averaged over sex. Prediction intervals incorporated the stochastic variation in the outcome data and the uncertainty in estimating the model coefficients [15].

RESULTS

A total of 484 subjects were recruited into the study, representing a 70% response rate. Of these, the 416 who reported past-year sexual activity were included in this analysis. Additionally, the 7% of person-months with a degree of zero among those active in the past year were dropped. Table 1 provides descriptive statistics for this subpopulation included in the analysis. Over half the population was female and below the age of 30 years. The average number of lifetime sexual partners was 4.58 (median = 3) and varied significantly by sex (although the difference in medians was smaller). Men also had higher past-year total partners and unprotected partners compared to women. Most of remaining person-time was monogamous, but there were clear differences in degree by sex: overall 12% of person-time was concurrent, with 21% of men's time concurrent compared to 3% of women's' time. Across sex, the mean was 4.77 total acts per month and 4.34 unprotected acts, indicating the majority of acts were unprotected. Finally, the estimated HIV-1 prevalence in the study population was 4.4%, with infection twice as high among women compared to men. No HIV-2 infection was observed.

In the first model for total acts, there was no evidence of coital dilution. As shown in Table 2, the incidence rate ratio (IRR) for total acts comparing person-months with a degree of 2 to months with a degree of 1 was 2.07 (95% CI = 1.85 – 2.33). For a degree of 3, the IRR was 2.93 (95% CI = 2.35 – 3.69). The conditional rate for males did not significantly differ from females. There was a negative trend for maximum partnership duration: the number of total acts declined by 1% with each increasing year of partnership duration. In the second model for unprotected acts, there was some dilution, but only for months with 3 concurrent partners. The IRR for unprotected acts comparing person-months with a degree of 2 to months with a degree of 1 was 1.94 (95% CI = 1.72 – 2.20). For a degree of 3, the IRR was 2.41 (95% CI = 1.90 – 3.11). Both sex and partnership duration were not significantly associated with unprotected acts in this model.

Figures 1 and 2 show the predicted number of acts by degree with 95% prediction intervals. In Figure 1, the predicted total and unprotected acts are plotted for males and females separately, averaging over partnership duration. The predicted number of acts increases with degree, but the slope over degree is more gradual for unprotected acts compared to total acts. Expressed as per-partnership rates, the females with 2 and 3 partnerships would have 4.3 and 4.1 acts per partnership compared to 4.1 among monogamous females; males with 2 and 3 partnerships would have 4.5 and 4.4 acts per partnership compared to 4.4 among monogamous males. The predictive uncertainty is higher with degree-3 months both because of the greater variance in outcomes and greater uncertainty in coefficient estimates. The predictions are non-overlapping when comparing degree-2 to degree-1 person-months for both total and unprotected acts, implying no evidence for complete dilution. Figure 2 plots the predicted total and unprotected acts by degree and partnership duration, limited to degree-1 and degree-2 person-months for clarity. In both panels, the predicted number of acts for degree-2 months is higher than degree-1 months

across all durations. The predicted total acts declines with higher duration but is flat for unprotected acts, consistent with the two model coefficients for degrees 1 and 2.

DISCUSSION

In this study, we observed no coital dilution in total acts for up to 3 concurrent partners, nor for unprotected acts with up to 2 concurrent partners. Dilution was only observed for unprotected acts in the context of three concurrent partners (degree > 2). But there were few of these multiple concurrency person-months (1.9% of total months), and the confidence interval for this estimate does not exclude 0% dilution. If the point estimate is correct, this implies a moderate increase in condom use for these uncommon, high-risk time periods.

Dilution research to date has investigated populations in eastern and southern Africa, with mixed results. Gaydosh et al. find some dilution among residents persons on Likoma Island in Malawi in which behavioral reporting of partnership behavior within couples was linked, but conflicting evidence among young Kenyans in an analysis from another study [5]. Although different statistical approaches were used across their analyses, dilution was always defined as a binary outcome: no sexual acts with a non-primary partner in the same interval over which degree is considered. Morris et al. found evidence for dilution in studies from Thailand, Uganda, and the United States using measures based on average coital frequency in the last 6 or 12 months, but differences in study design did not permit quantitative comparisons across the countries [6]. Delva et al. define dilution in terms of a weekly per-partner coital frequency in South Africa, using similar statistical modeling methods as ours, and find little evidence of coital dilution [8]. Delva's is also the only other study to distinguish between acts protected and unprotected by condoms (using a binary indicator for consistent vs. inconsistent/no use), and they found no difference in dilution by protection. In contrast, we found moderate unprotected dilution in months with multiple concurrencies (degree > 2).

The variability in results across these few studies on dilution to date may stem from several factors. One factor is differences in measurement methods, which makes comparisons difficult. While most of the studies have used some sort of event-history calendar to assess dilution, they have used quite different exposure and outcome measures. Dichotomizing dilution outcomes, in particular, leads to loss of information [16]. Therefore, we suggest that further empirical research on dilution adopt standardized measures for degree and acts as continuous variables, differentiating between total and unprotected acts.

Studies should also incorporate regression covariates useful for mathematical models of transmission dynamics. For example, we found that partnership duration is associated with coital frequency for long-term partnerships. By contrast, the effects of individual-level covariates, like age of the subject, are more complicated to assess because the attributes of both partners will influence their coital frequency, with the possibility of both main effects and interactions. It is

therefore necessary to carefully interpret associations on coital frequency for individual-level variables like age (or sex). Our models feature the subject's sex as an individual-level variable, but the purpose is to assess whether coital frequency is balanced for males and females, as it should be in a representative sample from a predominantly heterosexual population like ours (we did not limit survey questions to opposite-sex partners, but observed no same-sex partnerships). Along with standardized measurement, common statistical methods are needed. Statistical models like Poisson regression likely do not account for the overdispersed distribution of the coital frequency outcomes [8], so more flexible models like the negative binomial model should be considered.

Variability in dilution results across studies may also reflect actual differences in behavior across populations. For example, if the previous studies had measured dilution consistently, one could speculate that the dilution observed in populations with lower HIV prevalence (Malawi and Kenya) partially explains the disparity in disease burden between these areas and the higher-prevalence areas with no dilution (South Africa). But that hypothesis would be inconsistent with our results, since there was no dilution observed in our population, and it has the lowest disease burden. Again, however, while the goal of this body of research is to provide scientifically comparable results, the variation in measurement precludes such comparisons at this point.

All else equal, the protective effects of dilution on disease transmission will vary with the prevalence of concurrency, the rate of acts, the proportion of acts that are protected, and the effectiveness of condoms to prevent transmission. The probability of infection for a susceptible person given contact with an infected person will be a product of the two risks of transmission per contact, with and without condom use, given degree:

$$1 - \left[(1 - \tau_u)^{\alpha_{u,d}} \times (1 - \tau_p)^{\alpha_{p,d}}\right]$$

where τ is the probability of transmission per act, α is the number of acts, u connotes unprotected acts, p connotes protected acts, and d connotes degree. Therefore, dilution is defined as heterogeneity in α with respect to d. Because all of these parameters may vary across populations, along with other protective factors like male circumcision [17, 18], an ecological comparison of α_d across studies is insufficient to quantify the epidemiological impact of dilution.

Mathematical modeling of HIV transmission dynamics is one framework that overcomes this limitation. Sawyers et al. modeled the effects of dilution on HIV prevalence, where dilution was expressed as the reduction in the transmission risk in "non-primary partnerships" [1]. Their simulated population could engage in primary and non-primary partnerships, having an effective maximum degree of 2, but dilution applied to non-primary partners only. In their model, the designation of primary/non-primary partner is fixed at the time the partnership begins, so a non-primary partner remains "non-primary" even if the primary partnership ends. It is worth noting that this assumption implicitly links concurrency to reduced total coital frequency in the population, a pattern that is not empirically supported, and would bias the observed effects of concurrency

downward. Under these assumptions, they found that dilution above 25% resulted in disease extinction for concurrency levels up to a point prevalence of 14% (the maximum they tested). With our results translated to their metric, we effectively observed 0% dilution for total acts and 7% dilution for protected acts in this population where 10% of person-time was concurrent (with large variations in degree by sex). So even under their very conservative model assumptions, concurrency would still be likely to increase HIV transmission in Ghana. Our future mathematical modeling work will investigate the effects of dilution with more robust stochastic network models given the epidemiological parameters observed here.

Limitations. As noted, degree within each month corresponds to the number of active, overlapping partnerships within that month. We conservatively assumed that months in which one partnership ended and another started were not concurrent, but this may be a downward misclassification in degree. In a separate simulation study (not shown) we found that as long as any such misclassification is not correlated with coital frequency, the bias on IRR estimates would be conservative. A second limitation related to degree is degree truncation, due to the event history survey approach focusing on the last three partners in the prior year. This will also have a conservative impact, and requires modeling the relationship between degree and acts as a nonparametric function. However, the proportion of the population with a degree greater than 3 at any point in time is likely very small in both our target population. In the prior year, only 7% of men and 1% of women had 4 or more *cumulative* partners, and the degree would be much less than this. Another limitation is that the measurement of acts, total and unprotected, is subject to misreporting. This would only matter if misclassification was correlated with degree, leading to a conservative bias if those with higher degree were more likely to underreport their acts, and a positive bias otherwise. Finally, our modeling approach does not explicitly incorporate the temporal relationship between person-months: the time-effect was only controlled via the ever-increasing partnership duration for ongoing partnerships. More structured time-series approaches may better predict the relationship of interest but did not fit our data well.

Conclusions. Coital dilution has received little attention despite its potential to mitigate the effects of concurrency on HIV transmission dynamics. Our findings suggest that the impact of dilution in our population would be minimal. To verify this, we would need a well-specified mathematical model of HIV transmission dynamics that incorporates empirically observed differentials in degree, dilution given degree, condom use, and other biological and behavioral factors influencing transmission potential. Temporal exponential random graph models [19] is a promising statistical and mathematical framework for such analysis. These models will hopefully improve on past attempts to model dilution [1] to address important epidemiological questions on the impact of concurrency with dilution on population-level HIV transmission in sub-Saharan Africa.

ACKNOWLEDGEMENTS

The following provided invaluable assistance with this study: Kamil Fuseini, Fidelia Dake and the staff at the Regional Institute for Population Studies at the University of Ghana, the Ghana AIDS Commission, our field staff (Vincent Kantah, Patrick Nyarko, Charlotte Ofori, Cecilia Segbedji, Maame Yaa Konamah Siaw, Bilaal Tackie, Habakkuk Tarezina, Solomon Tetteh), and the laboratory staff at NMIMR (Prince Parbie and Joyce Appiah-Kubi). We also thank Martina Morris, Steven Goodreau, and other members of the UW Network Modeling Group for comments on this study.

TABLES

Table 1. Demographic, Behavioral and Epidemiologic Characteristics of Sexually Active Adults in Agbogbloshie, Accra, Ghana, 2012 (n = 416)

	Total		Males		Females	
	% / mean	*95% CI*	*% / mean*	*95% CI*	*% / mean*	*95% CI*
Sex						
Male	42.5%	36.6% – 48.5%				
Female	57.5%	51.5% – 63.4%				
Age						
18-29	55.7%	49.6% – 61.9%	50.5%	41.7% – 59.3%	59.6%	51.8% – 67.4%
30-39	29.1%	23.7% – 34.4%	33.0%	24.8% – 41.2%	26.2%	19.6% – 32.7%
40-49	15.2%	10.3% – 20.0%	16.5%	9.8% – 23.2%	14.2%	8.5% – 20.0%
Sexual History						
Total Partners, Lifetime	4.58	3.67 – 5.49	6.77	4.79 – 8.75	2.99	2.71 – 3.26
Total Partners, Past Year	1.48	1.33 – 1.63	1.90	1.57 – 2.22	1.17	1.09 – 1.25
Unprot. Partners, Past Year	1.20	1.11 – 1.29	1.37	1.20 – 1.53	1.07	0.98 – 1.16
Degree[1]						
1	89.1%	86.4% – 91.8%	79.3%	73.9% – 84.6%	96.3%	85.7% – 95.8%
2	9.0%	6.7% – 11.4%	17.1%	12.4% – 21.9%	3.0%	1.3% – 4.7%
3	1.9%	0.1% – 3.0%	3.6%	1.3% – 5.8%	0.1%	0.0% – 1.5%
Partnership Duration[1,2]						
Maximum Duration	5.9	6.2	5.3	5.3	6.4	6.8
Coital Frequency[1]						
Total	4.77	4.27 – 5.28	5.52	4.59 – 6.45	4.22	3.69 – 4.75
Unprotected	4.34	3.89 – 4.80	4.75	3.96 – 5.54	4.05	3.52 – 4.57
HIV-1 Infection						
Infected	4.4%	1.8% – 7.0%	2.7%	0.0% – 5.4%	5.6%	1.6% – 9.7%

[1] Degree, maximum partnership duration, and coital frequency were calculated on person-months, whereas the remaining variables are fixed and calculated on persons.
[2] Partnership duration defined as the length of the longest-running partnership in any person-month

Table 2. Incidence Rate Ratios and 95% Confidence Intervals for Monthly Total Acts and Unprotected Acts

	Total Acts			**Unprotected Acts**		
	IRR	*95% CI*	*p*	*IRR*	*95% CI*	*p*
Degree						
1	1.00			1.00		
2	2.07	1.85 – 2.33	<0.001	1.94	1.72 – 2.20	<0.001
3	2.93	2.35 – 3.69	<0.001	2.41	1.90 – 3.11	<0.001
Male	1.06	0.99 – 1.14	0.08	0.99	0.92 – 1.07	0.85
Duration	0.99	0.99 – 1.00	<0.01	1.00	0.99 – 1.00	0.51

FIGURES

Figure 1. Conceptual Examples of Coital Dilution. Under 0% dilution, a person has 4 acts per month with a main partner (circle), adds a secondary partner (square) also with 4 acts per month, and the per-partnership acts remains constant. Under 100% dilution, the act rate per partnership drops to 2 upon concurrency. Dilution does not require that acts be equally divided as they are here for illustration purposes.

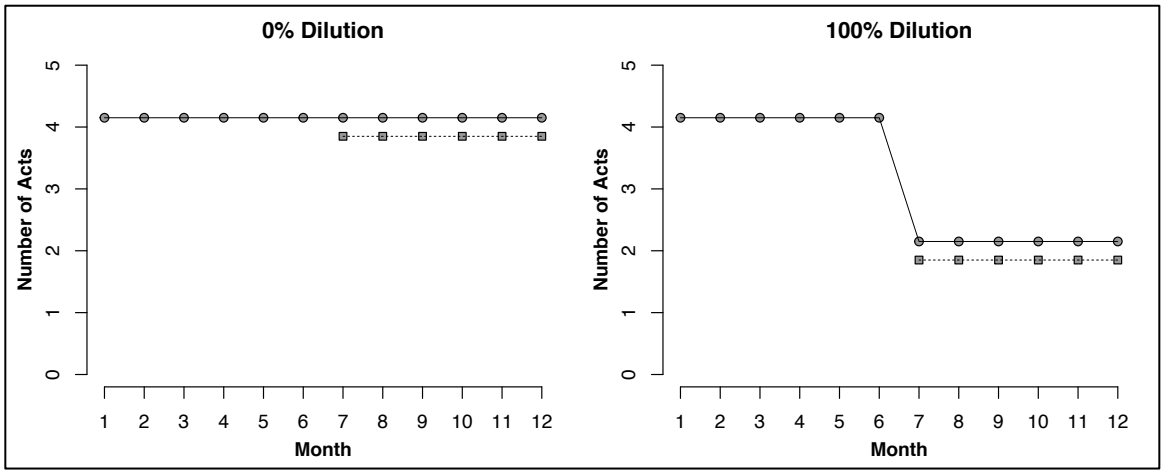


Figure 2. Predicted Number of Total Acts and Unprotected Acts by Monthly Degree and Sex among Sexually Active Adults in Agbogbloshie, Accra, Ghana.

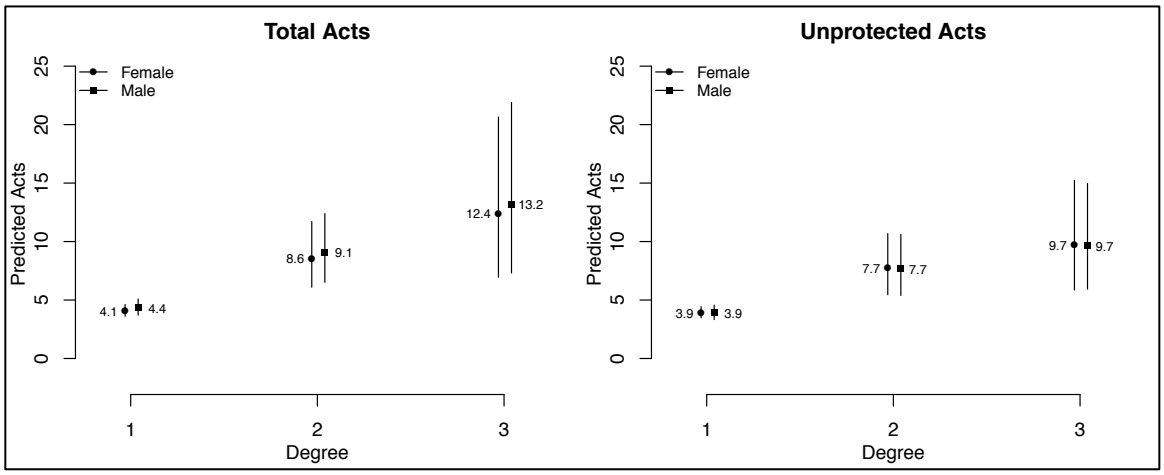
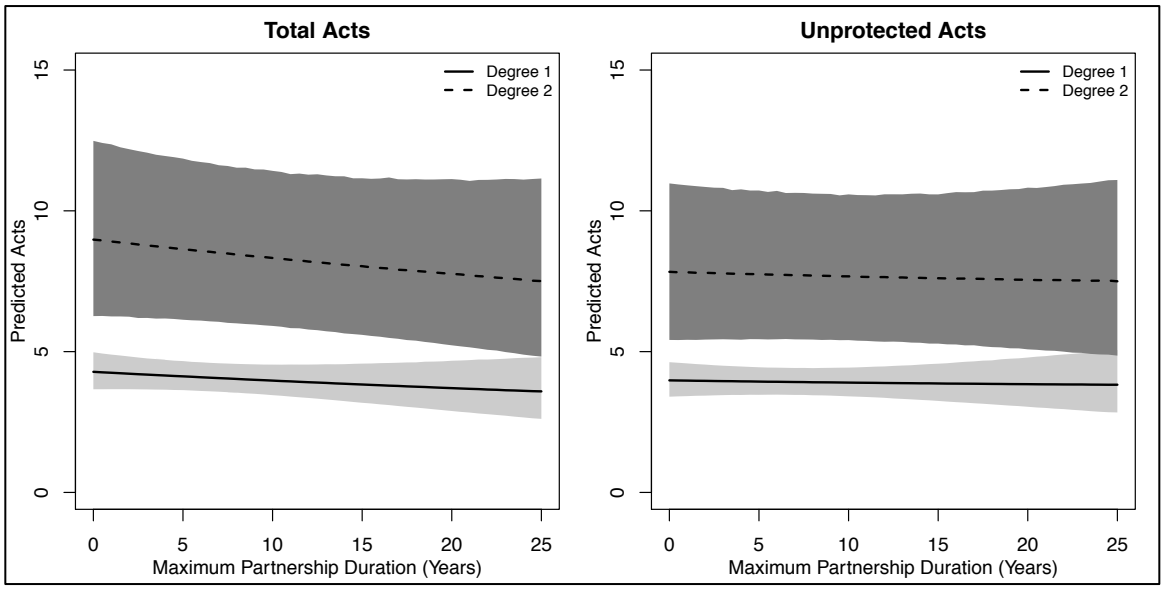


Figure 3. Predicted Number of Total Acts and Unprotected Acts by Monthly Degree and Maximum Partnership Duration among Sexually Active Adults in Agbogbloshie, Accra, Ghana. Degree was capped at two for visual clarity. Maximum partnership duration is the duration in years of the subject's longest-running partnership in any month.



Chapter 2: Effectiveness of Male Circumcision for HIV-1 Prevention Depends on Contact Network Structure

Effectiveness of Male Circumcision for HIV-1 Prevention Depends on Contact Network Structure

Samuel M. Jenness et al.

March 12, 2015

ABSTRACT

Background Targeting implementation of effective biomedical HIV prevention tools like male circumcision based on behavioral composition of the population is needed to maximize public health resources. Additionally, the design and rollout of combination prevention packages depends on the synergistic effects between biomedical tools and behavioral change. We investigate how dynamic sexual network structures influence the impact of circumcision among heterosexuals in Sub-Saharan Africa (SSA).

Methods With a dynamic stochastic network model for HIV-1 transmission dynamics, we parameterized a baseline model for western Africa using egocentric network data from a new population-based study in Ghana. Levels of male circumcision prior to sexual initiation were varied from 90% (baseline) to 10%. Levels of network connectivity were varied relative to the observed point prevalence of concurrency, from 50% to 150% of baseline levels.

Results The combined effects of circumcision and concurrency appears sufficient to explain the empirical population-level variation in disease burden across regions of SSA. For interventions, a 10% absolute change to concurrency prevalence is equivalent to an 80% change in circumcision prevalence. A 10% increase in circumcision would reduce HIV incidence 3.3-fold in low-concurrency settings and 1.1-fold in high-concurrency settings. Given current ART coverage, epidemic extinction only occurred during both large-scale changes to network connectivity and circumcision prevalence.

Conclusions Structures of dynamic sexual networks should be considered in targeting circumcision campaigns, and integrated into combination prevention packages. Synergistic biological and behavioral change in southern Africa could lead to a reduction of incidence to the low levels observed in western Africa.

1 INTRODUCTION

Combination packages for the prevention of HIV-1 in Sub-Saharan Africa depend on the synergy between individual interventions bundled together to achieve reductions in disease incidence [1]. Packages may combine effective biomedical tools like anti-retroviral therapy (ART) for prevention along with behavioral changes like increased condom use as secondary

components. Although no randomized clinical trials (RCTs) have found interventions for behavioral risk reduction alone effective in reducing HIV incidence [2], current research with combination packages seeks to understand how behavioral change components may complement biomedical tools to maximize their effectiveness [3].

Even for single-element interventions with no behavioral component, targeting epidemics based on the population-level composition may confer greater disease prevention benefits. Male circumcision is one technology with broad implementation [4] since three RCTs independently found a prevention benefit of over 50% [5-7]. Although a subsequent RCT that circumcised infected men was unable to quantify a risk reduction to uninfected female partners [8], mathematical models have suggested long-term indirect effects to women through population-level reduction in HIV incidence among men [9]. Ecologically, differences in circumcision have been suggested to explain the large regional disparity in HIV-1 burden across Sub-Saharan Africa (SSA) [10]. Male circumcision is nearly universal in many western African counties where disease prevalence is much lower than southern Africa, where circumcision is less common [11]. These disparities can help us understand the implications for wide-scale implementation of this intervention.

Mathematical modeling has been one approach to investigate the population-level impact of circumcision interventions, alone and in combined prevention packages [9]. These models can forecast the level of intervention coverage necessary to lead to widespread reductions in disease incidence, and possibly, epidemic extinction [12]. One early model predicted that circumcision interventions could lead to epidemic extinction if the circumcision coverage rose to 90% in tandem with large declines in the partnership acquisition rate (i.e., how quickly persons form new sexual partnerships) [13]. A more recent model suggested more modest declines with a prevention package incorporating treatment as prevention that required less extreme uptake [14]. An important contribution of that study was that major incidence declines would take much longer to see than typically studied in RCTs.

One limitation of both models, however, is that their mathematical representation of the sexual contact process is limited in its consistency with empirical data. These “deterministic compartmental” models for HIV require assumptions about contact rates that imply that the average person in the population has hundreds of partners over their sexual lifetime. (e.g., a mean 120 lifetime partners, with 40% of the population having greater than 145 partners [13]). Using more realistic values grounded in empirical data always leads to epidemic extinction, indicating a structural problem with the framework. Subsequently, models for combined biomedical and behavioral interventions that require unrealistic baseline levels of behavior also tend to require unrealistic behavioral interventions (e.g., a reduction of 40 lifetime partners [13]), limiting their utility in investigating novel prevention packages incorporating behavior.

In our study, we model dynamic sexual contact network structures based on observed data to understand the epidemiological synergy between networks and male circumcision interventions. This study investigates how changes to network connectivity, such as sex-structured age mixing along with multiple overlapping partnerships (concurrency), maximize the population-level effectiveness of circumcision to reduce HIV-1 incidence in SSA [15]. To this end, we address two related questions. First, does the varying prevalence of circumcision partially explain the regional disparities in the HIV-1 epidemic in SSA? Second, what are the synergistic effects of changes to the network structure and circumcision coverage, with respect to both targeting a circumcision-only intervention and also implementing a combined prevention package with both behavioral and biomedical components? To address these questions, we use a stochastic network for the HIV-1 transmission dynamics parameterized from sexual network data in West Africa.

2 METHODS

This study uses stochastic network models of HIV-1 transmission dynamics to simulate the conditions generating the HIV-1 epidemic in Accra, Ghana, and to test counterfactual conditions on modifiable biological and behavioral conditions influencing disease transmission. The simulated network structure follows the patterns of sexual partnership formation and dissolution over time. The full methodological framework for this study, including network model estimation, as well as epidemic model parameterization, simulation, and data analysis, are provided in the Supplementary Appendix **[LINK]**.

Parameters for the model are derived from an empirical study of sexually active adults in Accra, Ghana with the data collected specifically to inform these modeling activities [16]. Our Migration & HIV in Ghana (MHG) study was a 2012 cross-sectional study of adults in Agbogbloshie, an urban resource-poor area in Accra. A probability sample of the population was obtained, and study procedures included a standardized structured survey on demographics and sexual risk behaviors, as well as a diagnostic HIV-1/2 test. The survey provided egocentric-based network data on patterns of sexual partnership formation and dissolution. The study methods have been previously described [16] and further details are provided in the Supplement.

Dynamic Sexual Networks

Models for the sexual contact network within this epidemic model are based in the statistical framework of temporal exponential random graph models (TERGMs) [17]. TERGMs provide a statistical framework for estimating the generative parameters of sexual partnership formation and dissolution from sexual network data, and the resulting model can then be used to simulate

a dynamic network over time that maintains the targeted network statistics. An important strength of this framework is that the key targeted network statistics it needs can be obtained from easily collected cross sectional egocentric network data [18].

In this study, the dynamic network model included the following elements. This was a heterosexual transmission model, and therefore a network statistic was fixed such that men only partnered with women and vice versa. Several statistics were used to represent the momentary degree distribution, that is, the number of active, ongoing partnerships reported on the day of the study interview: Mean degree (the average number of ongoing partnerships per person at any time), and the prevalence of concurrent partnerships by sex (i.e., a degree > 1 on the day of interview). Both were modeled by sex: men had a higher mean degree and higher prevalence of concurrency from our data. Because the MHG data collection tool was limited to the last three partners in the prior year, the model included a similar constraint on the degree distribution. In addition, we modeled sex-structured age homophily. Typical of heterosexual mixing [19], persons in our study population tended to select partnerships similar in age (age homophily), but with a sex-based asymmetry wherein men were an average 5.38 years older than their female partners.

The relational dissolution model assumed a constant, homogeneous hazard that was based on the empirical distribution of partnership duration from MHG. We used Kaplan-Meier methods to estimate the survival distribution of reported partnerships based on their start, end dates, and right-censoring for ongoing partnerships. In the epidemic model, the partnership dissolution parameter was adjusted slightly to calibrate the simulated HIV prevalence to that observed in the empirical data. This involved adjusting the estimated mean duration from 3.97 to 3.08 years. This latter duration leads to a mean lifetime partner total of 10.1, which was still consistent with our estimates of 9.6 cumulative lifetime partners in our study population.

Disease Progression and Transmission

This component of the model uses a set of processes for HIV progression given the natural course of infection and anti-retroviral therapy (ART) treatment profiles. Persons progress through HIV stages with evolving CD4 and HIV viral loads, which are dependent over time. The CD4 progression model assigns base CD4 count upon infection conditional on sex, and the downward slope conditional of age at infection [20]. Baseline values and rates of decline are listed in Table S1 in the Supplement. HIV viral load follows a trajectory of peak viremia during acute-stage infection, followed by a viral set point during chronic infection, with a subsequent rise during late-stage infection onto AIDS and disease-induced mortality [21-24].

Upon infection, persons are randomly assigned a CD4 count for ART initiation, based on clinical data on medical care utilization in Ghana [25]. An upper limit on possible values was imposed that matched the current national AIDS policy of Ghana for initiation: persons did not

initiate ART until they fell below a CD4 count of 350. Additionally, there was an overall cap on ART coverage given estimates that only 30% of those eligible for ART have received it [26]. Upon ART initialization, persons were partitioned into full and partial adherence groups such that the average levels of HIV viral suppression matched broad empirical estimates for SSA [27]. ART adherence was associated with an increase in CD4 to pre-infection levels and a reduction in VL to suppression [25, 28, 29].

Disease was transmitted over active partnership dyads given the network model structure at any time step. The per-partnership transmission rate started with the Hughes [30] statistical model for the per-act transmission probability, which linked it to the HIV viral load of the infected partner, condom use, and the sex, age, and circumcision status of the susceptible partner. The final per-partnership transmission rate was a function of the per-act transmission probability and the number of acts per unit time. See Section 7 of the Supplement.

Baseline and Counterfactual Models

This study compares the endemic HIV prevalence and incidence as a function of two intervention-related behaviors: circumcision and concurrency. The baseline model starts with the disease prevalence observed in Ghana via the MHG study MHG study, representing a typical HIV-1 epidemic in a high-risk urban population in western Africa. to model the HIV-1 epidemic in a high-risk urban population in western Africa generally. As we vary the behavioral inputs, the patterns shift to the levels of circumcision and concurrency more commonly reported in southern Africa.

The baseline values for circumcision were 90% among males. Counterfactual models varied the prevalence of circumcision from this high level down to 10% in increments of 10% absolute proportion circumcised. Variations in circumcision prevalence are based on a childhood circumcision intervention, so all occur before entry into the sexual risk population. The targeted circumcision prevalence remains steady for the duration of each simulation.

Baseline concurrency varied strongly by sex, where 17.8% of men and 2.8% of women were categorized as concurrent on the day of study. The sex asymmetry is consistent with most SSA heterosexual populations [31], but levels of concurrency are known to vary considerably across the region [32], and to reflect some measurement error [33]. Therefore, the counterfactual models varied the levels of concurrency from 50% to 150% of that observed. This ranges captures much of the observed variation across the region. Like circumcision, counterfactuals on concurrency are modeled as stable conditions for the entirety of the simulation with no separate start time for changes.

In all scenarios, baseline and counterfactual, the mean degree and partnership duration do not change, so the per capita number of partnerships at any single time point, the partnership

acquisition rate, and the expected cumulative number of lifetime partners is the same in all scenarios.

A total of 99 scenarios were modeled in simulations: 11 levels of concurrency by 9 levels of circumcision. Each model was simulated 100 times over 100 years of simulation time to establish the endemic epidemic levels. Section 9 of the Supplement describes how this modeling approach differs from others that attempt to model the historic trajectory of the HIV epidemic from its outset with minimal to no empirical data on historical risk behavior. The results presented here are the summary of the final 100 time steps (~ 2 years) across all simulations within a parameter set. Given the stochasticity in the model, we present both the means and the interquartile ranges for the main outcomes.

3 RESULTS

The baseline model generated an endemic HIV prevalence of 4.7% (IQR = 3.7%, 5.6%), by design matching the level observed in the seroprevalence data in MHG. There was significant stochastic variability in prevalence and incidence, particularly for the lower-incidence transmission models like the baseline model. Figure 1 shows the time series for the predicted HIV prevalence across all simulations, with both the individual simulations and the mean over the set. The baseline model yielded an HIV incidence of 0.43 per 100 person-years (IQR = <0.00, 0.51).

Figure 2 summarizes the endemic HIV prevalence across all parameter sets, where prevalence rises from black to red. Concurrency is on a scale relative to the baseline model (100%), whereas circumcision prevalence is on the absolute scale. As expected, endemic prevalence was greater as concurrency rose and circumcision fell, both independently and together. Figure S1 in the Supplement shows the gradient across parameter sets for disease incidence; the trends were similar to endemic prevalence.

In Figure 3, we compare counterfactuals that independently vary concurrency and circumcision from baseline levels. The left panel holds circumcision fixed at the baseline level (90%) while varying concurrency, while the right panel holds concurrency fixed at the relative baseline level (100%) while varying circumcision. The smoothed lines provide the trends across counterfactuals. The steeper slope of the left panel indicates that reducing levels of concurrency yields larger reductions in prevalence than those obtained by increasing levels of circumcision. Indeed, the comparison of slopes underrepresents this difference since concurrency varies relatively while circumcision varies absolutely. In absolute terms, the range of concurrency for women is 1.4% (50% relative) to 4.2% (150% relative); for men it is 8.9% (50% relative) to 26.7% (150% relative). If men's absolute concurrency were varied by $\pm$10 percentage points, which corresponds to a $\pm$30% relative change, the disease prevalence ratio in the high

concurrency to low concurrency models is 8.2. This is greater than any 10% difference in circumcision across the gradient. Figures S4 and S5 in the Supplement show these comparisons assuming different levels of baseline concurrency and circumcision conditions.

Table 1 numerically compares disease prevalence and incidence for all levels of circumcision for three levels of relative concurrency: baseline (100%), 50% of baseline, and 150% of baseline. Given baseline concurrency, HIV prevalence would be 21.6% (IQR = 20.4%, 22.9%) if the proportion circumcised were 50%, and 30.6% (IQR = 29.4%, 31.8%) if it were 20%. Holding baseline levels of circumcision constant, the prevalence would be 0.2% (IQR = 0.0, 0.0) if concurrency were half of observed and 23.2% (IQR = 21.8%, 24.3%) if it were 150% of observed.

The means for several of the low-prevalence models fall outside their IQRs because the outcomes are not normally distributed. One contributing factor is epidemic extinction, a phenomenon that occurs as the system nears the threshold for sustainable transmission. Figure S2 shows the probability of epidemic extinction, defined as the prevalence reaching 0 infected persons in the population, across the parameter sets. In the lowest-risk set (relative concurrency = 50% and absolute circumcision = 90%), 28% of epidemics went extinct. Extinction was only observed in three of the 99 scenarios. This was partially a function of the number of simulations run; we would expect further extinctions in higher risk sets if more simulations were conducted. However it does accurately reflect the larger differences in the probability of extinction.

The incidence rate ratios (IRR) comparing varying levels of circumcision within levels of relative concurrency are also provided in Table 1. Under baseline concurrency, a 10% circumcision rate is associated with a 7.8-fold higher incidence rate compared to the baseline circumcision rate of 90%. In the lowest level concurrency scenario, the ratio is nearly 140-fold, while in the highest scenario that 80% change only yields 2-fold higher rate. Figure 4 shows this graphically for concurrency scenarios +/- 20% relative from baseline. Lower levels of concurrency are associated with steeper slopes for the IRR relating similar changes in circumcision prevalence. This indicates that for the same change in the prevalence of circumcision the effect on incidence would be greater in low concurrency settings. The slopes decline non-linearly with higher concurrency as the epidemic reaches a threshold level.

4 DISCUSSION

In this mathematical model for HIV-1 transmission dynamics, we investigated the relative contributions of network structure and male circumcision on HIV incidence in sexually active men and women in SSA. Our study builds on prior research by incorporating elements of the sexual network structure that allows a fit to observed HIV prevalence data with a set of empirically grounded parameters for sexual behavior that could be targeted for HIV combination

prevention interventions. This study provides the best estimates of the potential impact of combined circumcision and behavioral synergy paired with the newest empirical data on per-act transmission rates, natural HIV disease progression, and ART initialization and adherence in SSA.

Our primary empirical finding is that differences in circumcision alone do not explain the differences in HIV burden across these countries, but they are an important component along with network structure [11]. Circumcision and network structure may combine to help explain the disparities in HIV burden observed today [27], ranging from 1-2% in western Africa to 7-10% in eastern Africa to 15-25% in southern Africa. Mainly due to limitations in mathematical modeling methods, no models had answered this question this to date. Prior models that have investigated the effects of network structure alone, without circumcision, on HIV transmission dynamics used simplistic assumptions about demographic change [34], or did not incorporate HIV treatment [32]. Our model incorporates modern mechanics for HIV disease progression [20], entry into medical care [25], adherence to ART [27, 35], and the relationship between disease progression on transmission in long-term serodiscordant partnerships [30].

Compared to the otherwise low-level HIV-1 epidemics in western Africa, our empirical study population in Agbogbloshie represents a high-risk group: HIV-1 seroprevalence in MHG was over twice that estimated of in Ghana nationally (4.7% versus 1.9%). Our study provides evidence to suggest that one reason is the higher levels of concurrency relative to Ghana as a whole [16]. This may reflect the place of Agbogbloshie as a hub of circular migration. More modest levels of concurrency (e.g., 70% of that observed in MHG) are similar to national levels and reproduce the national epidemic levels holding circumcision prevalence constant [36]. With respect to international comparisons, our 50% relative concurrency counterfactuals are similar to those in rural Uganda, where the HIV prevalence is 6-10%, matching our model predictions given lower levels of circumcision there [37].

Model outcomes have implications for the design and targeting of combination HIV prevention packages that incorporate behavioral elements [1]. Whereas previous models exploring the synergism between circumcision and behavioral change have required large behavioral changes from unreasonable behavioral baselines (e.g., a reduction of 40 lifetime partners) to achieve significant reductions in HIV incidence [13], our model requires more modest changes to behavior related to network connectivity without any changes to the overall partnership acquisition rates. Furthermore, changes in concurrency have dramatic non-linear threshold effects on disease incidence compared to changes in circumcision. Under baseline conditions, a 10 percentage point drop in concurrency would yield a greater incidence reduction than an 80 percentage point increase in circumcision. The counterargument, of course, is that even small population-level changes to behavior over a sustained time period may be infeasible

compared to a one-time minor surgery [38]. Still, the variations in concurrency prevalence within Ghana, and even larger variations across the SSA region, suggest that alternative behavioral norms can become established. This study suggests that network structural interventions are promising in combination with circumcision interventions.

Finally, even in the absence of change related to network structure, targeting biomedical tools based on a fixed network structure should be considered. Our models suggest that an individual-level intervention like circumcision confers the greatest reduction in disease incidence in low-concurrency settings. The underlying reason is because the individual protection from circumcision cannot be sustained in a more rapidly circulating disease setting as occurs when there is high network connectivity [39]. Therefore, the prevention effectiveness of circumcision interventions would be maximized in low-concurrency settings like rural Uganda in which the individual-level protection may be sustained [37], whereas combination packages incorporating behavioral change would be considered elsewhere.

Limitations

One limitation of this study is the need to use calibration for the mean partnership duration to match observed HIV prevalence in our target population. As described in more detail in the Supplement, this partially arises from the fact that a single partnership duration parameter does not capture the heterogeneity in the duration distributions. Future work will expand this model to include richer detail in partnership types, which will incorporate this heterogeneity.

Second, in contrast to some models, we did not explicitly simulate changes to concurrency or circumcision as explicit interventions with a defined time horizon. Instead, when interpreted as interventions, the changes in behavior or biological should be considered fixed at sexual debut. The main reason was that we were interested in both the empirical question on regional disease disparities, reflecting the natural dynamics of disease, as well as counterfactuals interpreted as interventions. Therefore, the counterfactual comparisons as interventions should be considered the maximum effect of that change, reflecting the equilibrium settling point of an epidemic once the intervention has reached its final coverage threshold.

Finally, there is an issue of network boundary specification [40]. Our baseline models were based on data collected from a high-risk core in an otherwise low-level epidemic. Persons who were in our target population reported sexual partnerships with those outside of it, including persons older or younger than our age eligibility criteria and persons outside our geographic sampling area [41]. Our egocentrically based network statistics were as if it were a composite unbounded network. Preliminary analyses suggest little *demographic* differences in the reported partner population ineligible for the study, but differences in their risk behavior and other traits are unknown.

Conclusions

The primary contribution of the study was to assess the synergistic effects of network structure on circumcision, both as empirical patterns explaining the historical HIV burden across SSA, but also as intervention components in a combined HIV prevention package. Future empirical and modeling research is needed to understand the impact of such interventions targeted to specific geographies and age groups, which was outside the scope of our current model. Our models may also be relevant for different biomedical interventions for at-risk individuals that modify their individual-level risk, such as pre-exposure prophylaxis.

5 TABLES

Table 1. Endemic HIV-1 Prevalence, Incidence per 100 Person-Years, and Incidence Rate Ratios Comparing Circumcision Levels within Three Levels of Concurrency Relative to Baseline.

	Relative Concurrency[1]								
	50% (M = 8.9%; F = 1.4%)			100% (M = 17.8%; F = 2.8%)			150% (M = 26.7%; F = 4.2%)		
Circumcision[2]	*Prev.*	*Inc. / 100 PY*	*IRR[3]*	*Prev.*	*Inc. / 100 PY*	*IRR*	*Prev.*	*Inc. / 100 PY*	*IRR*
90%	0.2 (0.0, 0.0)	0.01 (0.0, 0.0)	1.0	4.7 (3.7, 5.6)	0.43 (0.0, 0.51)	1.0	23.2 (21.8, 24.3)	2.20 (1.36, 2.87)	1.0
80%	0.4 (0.1, 0.6)	0.04 (0.0, 0.0)	3.3	10.4 (8.8, 11.7)	0.96 (0.52, 1.53)	2.2	25.7 (24.5, 26.9)	2.50 (1.46, 3.42)	1.1
70%	0.8 (0.4, 1.1)	0.07 (0.0, 0.0)	5.6	14.0 (12.3, 15.5)	1.31 (0.58, 1.76)	3.1	29.1 (27.7, 30.5)	2.90 (1.64, 3.91)	1.3
60%	1.3 (0.8, 1.8)	0.11 (0.0, 0.0)	9.5	17.8 (16.5, 19.2)	1.70 (1.13, 2.43)	4.0	31.6 (30.2, 33.1)	3.13 (1.75, 4.21)	1.4
50%	3.6 (2.6, 4.5)	0.30 (0.0, 0.49)	25.6	21.6 (20.4, 22.9)	2.06 (1.29, 2.75)	4.8	34.5 (33.6, 35.6)	3.48 (1.94, 4.61)	1.6
40%	6.9 (5.9, 7.8)	0.61 (0.0, 1.02)	51.6	25.4 (24.5, 26.5)	2.48 (1.45, 3.40)	5.8	36.9 (35.9, 38.1)	3.78 (2.14, 4.97)	1.7
30%	11.3 (10.1, 12.3)	1.00 (0.54, 1.61)	84.6	28.0 (26.7, 29.4)	2.75 (1.58, 3.70)	6.4	39.3 (38.1, 40.6)	4.10 (2.28, 5.41)	1.9
20%	14.7 (13.7, 15.8)	1.34 (0.60, 1.82)	112.4	30.6 (29.4, 31.8)	3.03 (1.71, 4.10)	7.1	41.5 (40.5, 42.4)	4.39 (2.46, 5.81)	2.0
10%	18.1 (17.1, 19.0)	1.66 (0.67, 2.42)	139.6	33.1 (32.2, 34.1)	3.33 (1.85, 4.41)	7.8	43.5 (42.5, 44.4)	4.70 (2.66, 6.23)	2.1

[1] Concurrency is varied by proportions relative to the baseline values (100%), where the frequency varies by sex for males and females. The relative proportions are shown, then the absolute proportions for Males and Females.

[2] Circumcision is the absolute proportion of males circumcised upon entry into the sexual adulthood, where 90% is the base value in Ghana.

[3] IRR = incidence rate ratio comparing incidence rates by circumcision level within levels of relative concurrency.

6 FIGURES

Figure 1. HIV prevalence over burn-in and analysis time periods for the baseline Ghana model. The burn-in period lasts for the first 2,600 time steps (50 years). Endemic HIV prevalence level fit to seroprevalence data of 4.7%. Blue lines show results of each of 100 model simulations, and red line shows the mean across all simulations. Model results were analyzed by taking the mean of the final 100 time steps of each of the 99 simulation sets.

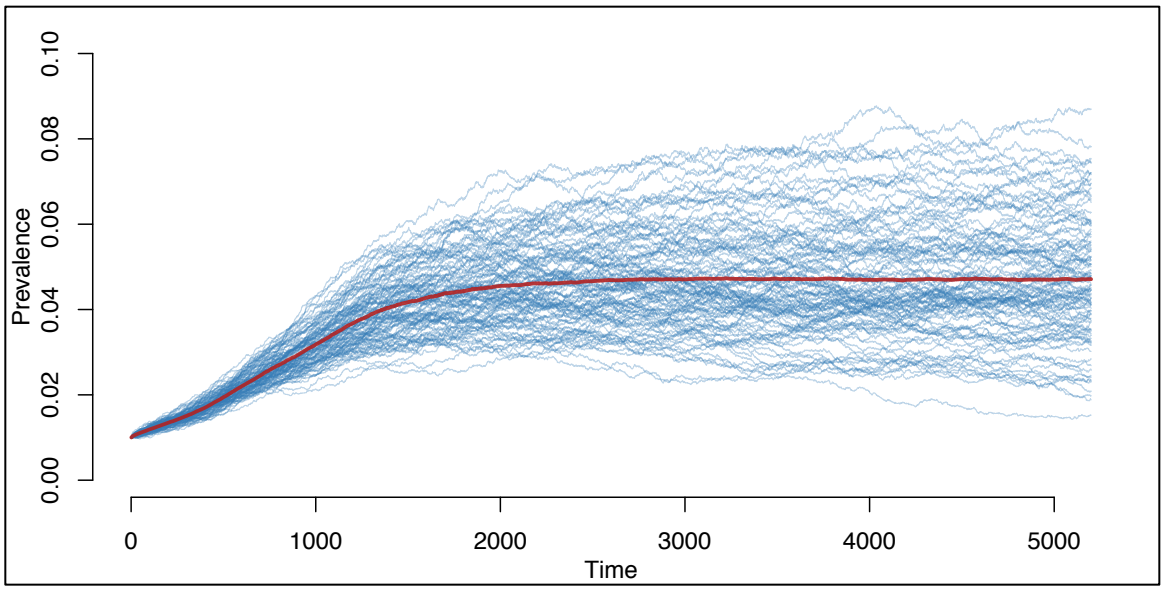


Figure 2. Endemic HIV prevalence by relative concurrency prevalence (50% to 150% of Baseline Values) and absolute circumcision prevalence (10% to 90% of men). Baseline model is 100% relative concurrency and 90% circumcision. Brightest red indicates highest prevalence and dark grey lowest prevalence.

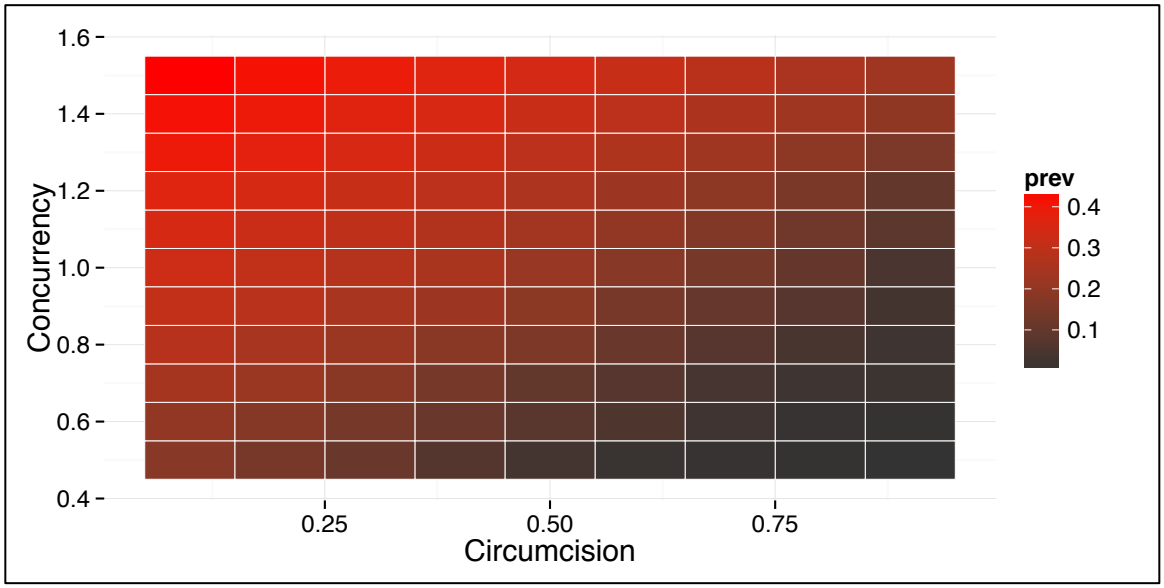


Figure 3. Distributions of endemic HIV prevalence comparing differences in relative prevalence of concurrency (50% to 150% of observed values) given baseline circumcision (left) and absolute prevalence of circumcision (90% to 10%) given baseline concurrency (right). Boxplots show the distribution of outcomes for the analysis period for each of the counterfactual models. Red lines are lowess smoothed curves over the means of across concurrency levels.

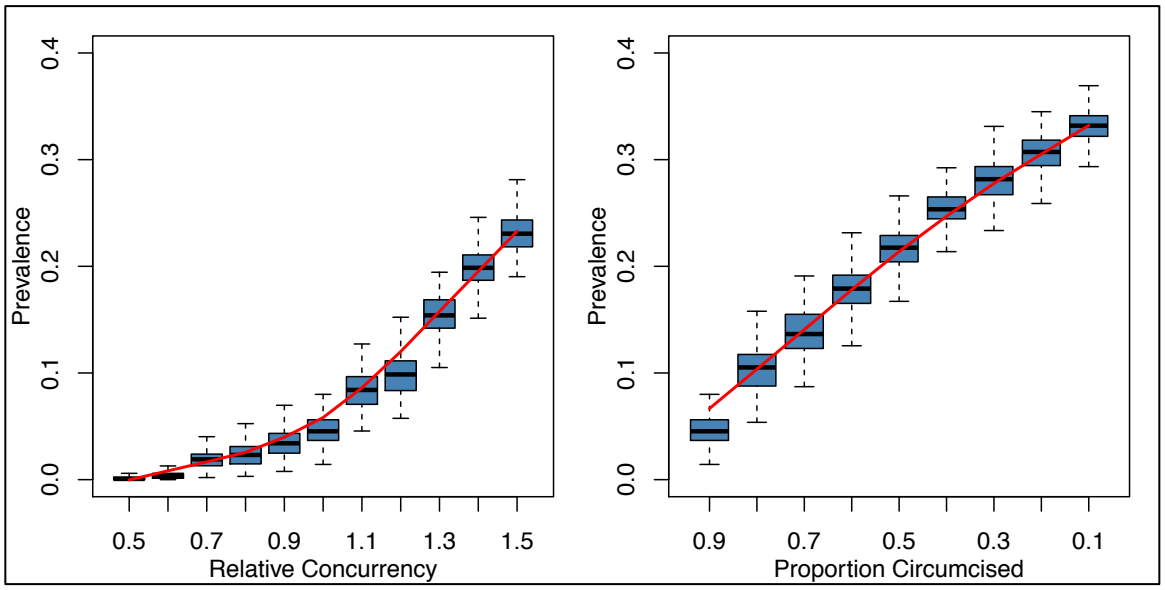
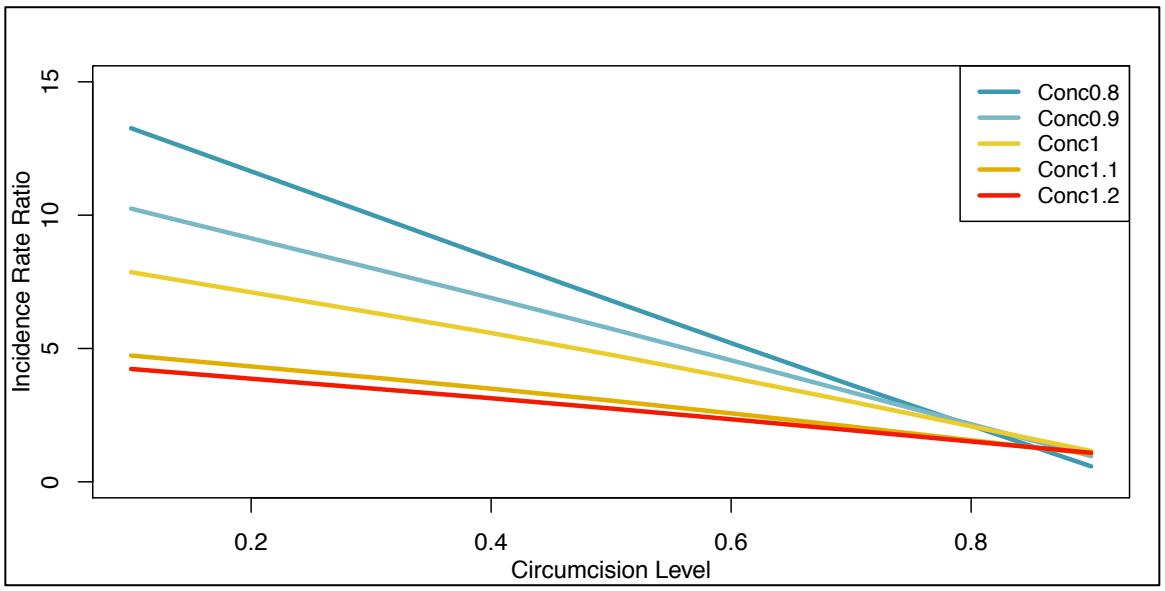


Figure 4. Incidence rate ratios comparing the relative increase in incidence by absolute circumcision prevalence for five sets of relative concurrency levels, from 80% to 120% of baseline values. IRRs were smoothed with a lowess smoother across all simulations to remove the effects of stochasticity. The steeper sloped lines for lower prevalence of concurrency show a greater relative benefit for a circumcision intervention.



Effectiveness of Male Circumcision for HIV-1 Prevention Depends on Network Structure

Supplementary Appendix

Samuel M. Jenness et al.
March 12, 2015

1 INTRODUCTION

This supplementary appendix describes the mathematical model structure of the accompanying paper in more detail, and also provides some supplement figures that are not included in the primary paper. Given our study's use of primary data to parameterize our models, we also provide further detail on the behavioral data underlying the network model.

The models presented in this paper are individual-level microsimulation models in which uniquely identifiable sexual partnership dyads are simulated and tracked over time. This partnership structure is made possible through the use of temporal exponential-family random graph models (ERGMs), which are described in detail below.

Along with this dynamic network simulation, the larger epidemic model includes demography (entries, exits, and aging), inter-host epidemiology (disease transmission), intra-host epidemiology (disease progression), and clinical epidemiology (disease diagnosis, treatment initiation, and adherence). Individual attributes related to these processes are stored and updated in discrete time over the series of time steps for the simulation.

Computationally, this model was programmed in R and C++ using *EpiModel* (www.epimodel.org), which is software developed by the authors for the purpose of simulating complex network-based mathematical models of infectious diseases, primarily HIV and other sexually transmitted infections. *EpiModel* is part of and based on *Statnet*, which is a suite of software in R that provides tools for the representation, visualization, and statistical analysis of network data [1].

This current model extends the general *EpiModel* disease platform to incorporate HIV-specific elements not part of the core software tools. It will be released as a stand-alone R package, *EpiModelHIV*, on both Github (LINK) and CRAN (LINK).

2 PRIMARY DATA

Unlike many mathematical models for HIV, this uses rich socio-behavioral data on dynamic contact network structures that were collected specifically for the parameterization of these models [2]. Our empirical study that formed the basis of model parameters was called the *Migration & HIV in Ghana (MHG)* study. MHG was a cross-sectional study of sexually active

adults in Agbogbloshie, Ghana in 2012. Agbogbloshie is an urban resource-poor area in the capital city of Accra, selected based on its hypothesized high-risk profile and lack of prior epidemiological research.

2.1 Design

The empirical study design has been described in detail in prior publications [3]. Briefly, MHG used a two-stage cluster randomized sampling scheme to obtain a probability sample of the population. Starting with an area census, we first randomly selected households with probability proportional to household size, and then randomly selected one adult household member. Given differences in household size, a weighting scheme was employed to account for differential inclusion probabilities. Eligibility criteria to participate were current residence in the selected household, age 18 to 49 years, and lifetime history of consensual sexual intercourse. The Institutional Review Boards of the University of Washington and University of Ghana approved all study procedures.

2.2 Measures

Trained field staff administered a standardized structured survey and drew serum via finger stick for a diagnostic HIV-1/2 test. The survey focused on demographics, migration and travel, and sexual behavior. For sexual behavior, summary data were collected on the number of lifetime sexual partners, past-year partners, and past-year partners with whom condoms were not always used. We used an event-history calendar to collect detailed partnership data for partners in the past year (up to three), with responses for each month during that period [4]. For each partner, data included the duration of the partnership and monthly information on the number of total and unprotected sexual acts.

For HIV testing, dried blood spots were collected in the field on standard filter paper and maintained in refrigerators before delivery to the Department of Virology, Noguchi Memorial Institute for Medical Research, University of Ghana for processing. Serum was tested on the INNO-LIATM HIV-1/2 test platform (Innogenetics, Belgium), shown to have good sensitivity and specificity for diagnosis and HIV type differentiation. All subjects who tested were asked to return in one week to receive their test results, which were provided by a trained nurse counselor. HIV-infected subjects were referred to medical care.

3 CONTACT NETWORK MODEL

At the center of this model is the dynamic network structure, which concerns the HIV-related sexual contacts that individuals make over time. A dynamic model for the formation and dissolution of sexual partnerships based on the empirical data above was formed with the statistical methods of exponential-family random graph models (ERGMs) [5]. ERGMs provide a

foundation for a statistically principled simulation of local and global network structure given a set of target statistics from empirical data. In this section, we describe the statistical and mathematical framework of this approach, and then discuss the unique parameterization of this model.

3.1 Mathematical Framework

Temporal ERGMs consist of a formation model and a dissolution model. The formation model governs how two actors not previously connected within the network pair, and the dissolution model influences how two currently paired actors dissolve their dyad. In the formation model, the probability of observing a set of *Y* relations among *n* actors given a set of attributes is expressed as:

$$\mathrm{P}(\mathrm{Y} = \mathrm{y}|\mathrm{n}, \mathrm{X}) = \exp\left(\frac{\theta' \mathrm{z}(\mathrm{y})}{\kappa(\theta, \mathrm{n}, \mathrm{X})}\right)$$

where $\mathrm{z}(\mathrm{y})$ are a set of network statistics, and θ are coefficients to be estimated from the model. The denominator of the equation is a normalizing constant, which given the potential number of network configurations of a sufficient size, is not estimable. Therefore, simulation-based estimation methods using Markov-Chain Monte-Carlo (MCMC) sampling are used to obtain maximum likelihood estimations for θ given the set of network statistics, z [6].

Network statistics may be of arbitrary complexity, and fall into two main classes: dyadic independent and dyadic dependent statistics [5]. Independent statistics are those in which the probability of dyad formation between any two nodes, i and j, does not depend on the existence of ties from i and j to other nodes (i.e., persons) in the network. Examples include the propensity for mixing by sex (i.e., heterosexual mixing) or age-based homophily (i.e., choosing a closely aged partner). In contrast, dependent statistics are those in which the probability of the tie between i and j depends on the presence of any ties from those two nodes to other nodes in the network. A common example is degree, which is the number of active ongoing ties for each node. The probability of forming a tie with a new potential partner may be hypothesized to depend on ties already in existence.

Dynamic network models also include a dissolution component that predicts the network statistics associated with ties dissolving conditional on their existence [7]. Within each time step, the formation and dissolution models are said to be separable: for any node the process of tie formation is independent of tie dissolution within a time step, but dependent over time. The implications of this assumption are that the size of the time step must be minimized to reduce the effects of the within-step independence.

The dissolution model may be estimated a number of ways. In this current study, we use the Edges Dissolution approximation method of Carnegie et al [8], in which a static ERGM was fit

and the formation coefficients statistically adjusted to account for edge dissolution. This method was chosen for computational tractability compared to fitting full temporal ERGMs. The two methods are implemented in the EpiModel software.

With this method, a complete network need not be observed for the purposes of dynamic simulation. We used egocentrically collected data, via the MHG Study, to query localized network measures (e.g., attribute-based homophily and degree). The MCMC-based simulation used for network model estimation is also used for the simulation of complete networks with features stochastically varying around target statistics from the data. The specific parameterization is provided below, but the egocentric framework is discussed in more detail elsewhere [9].

Finally, for the dynamic simulations two adjustment methods were used to account for the changes to the population composition over time. The first adjustment was to preserve mean degree (i.e., the average number of partners per person at any time) given fluctuations in population size with births and deaths. The offset method of Krivitsky and colleagues [9] holds this mean degree constant, a reasonable assumption for HIV/STI epidemic models. This adjustment also allowed the data from the original MHG study sample to be projected up to a network size representing the size of the underlying population: we started all simulations with a network size of 10,000 nodes, which grew over time with demographic trends.

The second coefficient adjustment method was to account for the presence of deaths or exits from the network as an exogenous competing risk for partnership dissolution, on top of the endogenous rates of partnership dissolution observed in the data. Since the empirical data were cross-sectional, we did not observe mortality rates to be able to partition the partnership dissolution into these two components. Therefore, we used an adjustment of the dissolution coefficients whereby the probability of tie persistence (the complement to tie dissolution) was adjusted upward to account for the competing risk from mortality.

3.2 Parameterization

The contact network model was parameterized from originally collected data in the MHG study described above. This section provides information on the structure of the formation and dissolution models, how the data were estimated, and methods for model calibration.

Formation Model. Given the aims of this study our network models were formulated to capture the unique sex and age-based mixing structure and degree distributions in a heterosexual population in West Africa. The network statistics were directly estimable via egocentric inference from the empirical cross-sectional data in MHG. The formation model included the following elements:

1. **Mean degree.** The overall number of dyads observed in the cross-sectional data was modeled with an `edges` term in the ERGM. This term is commonly used like an intercept in a generalized linear model, so that the remaining coefficients are interpreted in its base value. The number of edges in the model was based on a calculation of cross-sectional mean degree at the day of the study. As noted, the coefficient was adjusted to account for differences in the observed study sample and the simulated network size, preserving the mean degree. In the data, the observed mean degree was 0.91 overall, corresponding to an expected 4,544 edges in a network of size 10,000.
2. **Heterosexual mixing.** This model is designed to study purely dissortative mixing between women and men in the population (i.e., only heterosexual HIV spread). This was both a simplifying assumption of the model, and also reflecting the fact that no same-sex partnerships were observed in MHG. At the time of the study, same-sex activity was illegal in Ghana [REF], so there was strong bias against study participants reporting on this behavior. In the model, we parameterized heterosexual mixing as an offset term for the `nodematch` network statistic for the attribute of sex. This involved fixing the coefficient at a value of negative infinity so that the term was not estimated in the model fitting procedures. Simulations from the model fit were consistent with this constraint.
3. **Degree distribution by sex.** Without degree terms in the network model, the degree distribution (the number of persons with a momentary degree of 0, 1, 2, and so on) would follow a Poisson distribution with the rate parameter a function of the mean degree. In our data, we observed many fewer concurrent nodes (persons with a degree of 2+) than expected under this null model, and furthermore, the degree varied largely by sex: concurrency was much more common in men than in women. To capture this we included terms for the number of nodes with concurrency, differential by sex of the node. The baseline values for this were 17.8% for men and 2.8% for women. These were transformed into a number of concurrent nodes given the population sizes of men and women, and modeled using the `concurrent(by = "male")` statistic in ERGM.
4. **Degree constraint.** As described in prior reports [3], our data collection tool was capped at the last three partners in the past year, primarily to reduce response fatigue. This would miss a certain proportion of respondents who had a high momentary degree (4+). However, the *cumulative* number of partners was greater than 3 in only 7% of men and <1% of women; the momentary degree would typically be much less than even this. This constraint in the data collection influences the calculation of the mean degree and other degree statistics. Therefore, we used a constraint term in the model to cap the number of ongoing partners at any one time at 3.

5. **Sex-structured age homophily.** Typical with heterosexual mixing [10], our West African population tended to select sexual partners similar in age (age homophily), but with a sex-structured directionality. This directionality was such that men were on average 5.38 years older than women. To model this, age was included as a continuous attribute with values following the empirical data. A new ERGM term was coded using the `ergm.userterms` software package to model the absolute difference in ages with sex-structured directionality [11]. This network statistic includes both an offset value (5.38 years) and a residual sum after the offset that controls the variance of the distribution. Model diagnostics were checked to ensure that the offset mean and variance were preserved over changing age distributions that reflected demographic trends.

Dissolution model. In contrast to the formation model, the dissolution model includes only one network statistic, which is a function of the average duration of all partnerships. There have been several methods used to parameterize this statistic [12]. In any approach, the goal is to estimate a coefficient for the dissolution model prior to estimating the formation model. It is therefore entered as a fixed offset term that specifies the average duration.

In our approach, we use Kaplan-Meier methods to estimate the survival distribution of the empirical partnership durations from MHG [13]. The start and end dates of up to the three last partnerships in the prior year were recorded, by calendar month and year. To obtain daily intervals, we sampled from days of the month with a uniform distribution to obtain an imputed starting and ending date. All partnerships that were qualitatively described as "one-offs" by the study participant were coded as one day in duration.

Partnerships that had not ended on the date of the study were right-censored, whereas the duration for the remaining partnerships was considered fully observed. The KM survival curve was available down to the 68th percentile (only 32% of queried partnerships had dissolved by the study date). The network model dissolution coefficient is expressed as a λ coefficient in a geometric distribution. We estimated λ for a geometric distribution that matched the 75% percentile of the observed KM survival curve. The 75th percentile was 1,978 days, which matched a λ parameter of 1/1453 days. Therefore, the average partnership duration based on the empirical data was estimated at 1,453 days or 3.97 years.

This duration coefficient was used as the *only* calibration parameter in our epidemic model. In order to fit the model results to HIV prevalence observed in our diagnostic data, the coefficient was adjusted downward. With no adjustment, the endemic HIV prevalence was too low relative to observed data, and therefore we adjusted the mean of 1,453 days to 1,124 days (or 3.08 years) to calibrate. This 22% reduction is justified based on the comparison of the observed duration data to the homogenous geometric distribution probability mass function of

the original estimated mean duration. The geometric distribution with that parameter under-estimates the number short-duration partnerships and over-estimates the number of long-duration partnerships. For example, 28.6% and 9.6% of observed partnerships were under 1 years and 1 month of duration, respectively. The expected quantiles given the λ above are 22.3% and 2.0%, respectively.

Additionally, this 3.08 average partnership duration reflects an average number of lifetime sexual partners of 10.1 given a sexual lifetime of 37 years, which corresponds to the range of age minimum to maximum in our model (18 to 55 years). This closely reflects the best estimation of the average number of lifetime partners based on cumulative partnership data we observed in MHG (9.6 lifetime partners).

4 DEMOGRAPHY

In this model, there are three main demographic processes: entries, exits, and aging. Entries and exits are with respect to the sexually active population at which point persons become at risk of infection via heterosexual transmission, and so are modeled as starting at an age after birth and potentially ending at an age before death.

4.1 Entry/Birth

All persons enter the network at age 18, which was the lower age bound of the MHG study student eligibility criteria [3]. The number of new entries at each time step is a proportional function of the current population size at that time step. This entry rate parameter was fixed so that there was a stable average growth rate in the population to maintain the stationary population distributions observed empirically. The number of entries into the network at each time step was simulated by drawing from a Poisson distribution with the rate parameter equal to the fixed entry rate times the current population size.

Incoming nodes were randomly assigned a sex and, for males, a circumcision status, among other relevant attributes. The probability of assignment as male sex was based on a draw from a binomial distribution with probability parameter set to maintain the sex ratio distribution at the outset of the simulation. This initial sex distribution was 54.9% female and 45.1% male, which was estimated from the MHG data. The sex ratio imbalance likely arises from the high circular migration of men from the community [14], among other factors. The specified level of circumcision in the baseline model was that observed among men in the MHG study, 90%, similar to other West African countries [15]. The circumcision status was also randomly assigned at entry with by drawing from a binomial distribution with probability set to the specified level. The counterfactual models for this study varied that probability for incoming nodes.

4.2 Exit/Mortality

All persons exit the network by age 55, either from death or cessation of sexual activity. The upper limit of age 55 was enforced deterministically for everyone in the network. The exits due to death were stochastic processes based on natural (non-HIV) and HIV-related causes before that age. Background mortality was modeled with age-specific and sex-specific mortality rates derived from the World Health Organization's demographic life tables for Ghana [16].

Life table data provide the probability of death by age and sex for the population for the year 2010, closest to the date of our MHG empirical study. The mortality rates were then applied to active persons within the network through a stochastic process by drawing from a binomial series for each eligible person with a probability corresponding to that person's age and sex. Disease-related mortality was modeled based on clinical disease progression, and is described in Section 5 below.

4.3 Aging

The aging process in the population was linear in time for all active nodes. Nodes who exited the network from death were no longer active and their attributes were no longer updated. The unit of time step in these simulations was one week, and therefore, nodes were aged in weekly steps between the minimum and maximum ages allowed (18 and 55 years old). Whereas the eligibility criteria of the MHG study was between 18 and 49 years, we used an expanded upper age limit for the simulation to account for observed age-mixing between study participants and older persons within this 55 year age limit in the underlying study target population.

5 INTRAHOST EPIDEMIOLOGY

Intrahost epidemiology includes those processes related to disease progression within infected persons. The two main components of progression within the HIV context are CD4 count and HIV viral load [17]. Whereas the behavioral and demographic parameters were derived from the MHG study and related Ghana data, these clinical data come from large external cohort studies and clinical trials, referenced below.

5.1 CD4 Progression

Persons have a CD4 count assigned at infection. Post-infection, that count declines naturally in the absence of anti-retroviral therapy (ART) and rises with successful ART. The underlying model for the decline follows Pantazis and colleagues, summarizing a meta-cohort of Africans [18]. This model has a non-linear decline in CD4 count based on age at infection and sex within Sub-Saharan Africa. Women have a higher baseline CD4 compared to men, and the rate of decline is faster when persons are infected at an older age. The decline is non-linear as the slopes are calculated on the scale of square roots of CD4 count.

Table S1 shows example values of the base CD4 count at infection and then the years it takes to reach certain threshold values of 350, 200, and 100 cells. As noted, women have a higher base value and progress to AIDS more slowly than men. Persons infected at younger ages will progress more slowly than those infected at older ages across sex.

Table S1. CD4 Clinical Model

	Base Value	Years to Threshold		
	Base CD4	*350*	*200*	*100*
Males				
Age Infected				
25	518.4	3.53	7.50	11.10
35	518.4	3.24	6.90	10.22
45	518.4	2.73	5.79	8.57
Females				
Age Infected				
25	570.3	4.50	8.47	12.07
35	570.3	4.13	7.79	11.10
45	570.3	3.47	6.54	9.31

Disease-induced mortality is a stochastic process when the CD4 crosses a late-stage AIDS threshold. Similar to Eaton and Hallett [19], we model this threshold level at a count of 50 CD4. Once persons hit that threshold, disease-related death occurs on average within one year. This is parameterized by a series of binomial draws with probability equal to expected mortality in one year. Since this is a stochastic process of geometrically distributed waiting times, persons may not randomly die before they reach a CD4 of 0, but that lower limited is enforced as a deterministic process.

5.2 HIV Viral Load

Following the prior approaches [12], we parameterize changes in HIV viral load to account for the heighted viremia during acute-stage infection [20], viral set point during the long chronic stage of infection, and subsequent rise of VL at clinical AIDS towards disease-related mortality. A starting viral load of 0 is assigned to all persons upon infection. From there, the natural viral load curve is fit with the following parameters.

Post-infection, it takes 14 days to reach peak viremia [21], at a level of 6.7 $\log_{10}$ [21, 22]. From peak viremia, it takes 107 days to reach viral set point [23], which is at a 4.5 $\log_{10}$ level [21, 22]. The total time of acute to chronic-stage infection will depend on CD4 progression using the model specified above. Rise of viral load begins during AIDS-stage infection, which occurs when the CD4 reaches a threshold of 200. This again follows the Pantazis disease progress model

[18]. For men infected at age 45, that will occur 5.8 years after infection; for women infected at age 25, it will occur after 8.5 years.

During this late-stage infection period, viral load will rise from the set point of 4.5 $\log_{10}$ to a maximum of 7 $\log_{10}$ [17]. The time it takes to reach this maximum is also a function of the CD4 slope, and with the time it takes to transition from a CD4 of 200 to a CD4 of 50, at which point late-stage AIDS with a high probability of mortality will occur [19]. These transitions are deterministic for all individuals. The VL trajectory is for ART-naive persons. The influence of ART on this trajectory is described below.

6 CLINICAL EPIDEMIOLOGY

Clinical epidemiological processes refer to disease diagnosis and treatment. In our model, diagnosis and treatment occur simultaneously based on the CD4 level of the respondent. This follows the empirical research of Collini et al [24] in Ghana that provides the mean and standard deviations of CD4 upon treatment initiation, which were 120 and 88 respectively. This assumption that there is no waiting time between diagnosis and treatment implies that individuals do not change their sexual risk behavior between the two events; there are no representative data on this for Ghana.

6.1 Anti-Retroviral Therapy (ART) Initiation

Upon infection, each person is assigned a CD4 count at which they will initiate ART. This is drawn from a negative binomial distribution with mean and standard distribution parameters specified as stated above. Based on the statistical tests used, that study implied that the values were normally distributed. However, it was necessary to re-express this as distributed as a negative binomial because the parameter values (120 mean; 88 SD) result in 9% of the probability density below 0. We further scaled the standard deviation parameter for the negative binomial to 40 because the original value resulted in a tail where a considerable portion of the mass was above a threshold initiation value.

That threshold value for Ghana was based on the country's national AIDS policy, in which persons may begin treatment at or below a CD4 of 350. Therefore, the draw for the individual-level count of when to initiate ART was capped at this 350 CD4 count value. Individuals could not initiate therapy above this count, and the average count at ART initialization followed the mean parameters.

There was an overall treatment coverage enforced among those reaching below their threshold values. Treatment coverage was expressed as the proportion of treatment-eligible persons (those with a CD4 under 350) who had ever initiated treatment. While the data on this for Ghana are not widely available, one report suggests that this current level is 30%, and we

used this as the base parameter in the model [25]. Persons were allowed to initiate treatment, therefore, only if their current CD4 level had reach their targeted initiation value and less than 30% of eligible persons were currently “on ART,” even if their adherence to ART had lapsed.

6.2 ART Adherence

Adherence to ART was modeled to achieve an overall level of viral suppression among those who had ever initiated ART. Since ART adherence over the disease history is a complex dynamic process, with much uncertainty in how persons adhere to ART across disease stages, we used a calibration method in which we targeted that 80% of those who had initiated ART were virally suppressed. This conforms a recent UNAIDS analysis of viral suppression levels across Sub-Saharan Africa [26].

Modeling adherence in this way involved partitioning the on-treatment population into two groups: full adherers and partial adherers. Based on a meta-analysis of ART adherence that included many Sub-Saharan African countries, we set the full adherence proportion to 76% of the on-ART population [27]. Upon ART initiation, persons were assigned to the full versus partial adherence groups using a series of binomial draws with probability parameters set to this 76% value.

Full adherers remained on ART for the duration of their disease with no lapses. Sustained ART use conferred two clinical benefits: increased CD4 and decreased HIV viral load. Based on data from Collini from Ghana [24], the rebound in CD4 was modeled as a linear daily increase in CD4 level, with the slope conditional on sex of the person. CD4 count for males increased at an average of 9.75 per month; for females it was 11.6 CD4 cells per month. For full adherers, the CD4 increased back up to the level at initial infection, where it stayed for the duration of disease. This implied no disease-related mortality for those fully adherent on ART. ART was associated with a sustained decrease in HIV viral load within three months of initiation [28, 29], from the person's current level to viral suppression, which we specified at 1.5 $\log_{10}$, just below the limits of detection (1.65 $\log_{10}$) [12, 30].

Partial adherers cycled back and forth between ART use and non-use as a Markov process at each time step. The probability of cycling was calibrated to 50% at each time step, which resulted in the overall level of viral suppression of 80% among those who had initiated ART [26]. As a simplifying assumption and in the absence of robust data on rebound rates in western Africa, cycling off treatment resulted the reverse immunological and virological phenomena as cycling on treatment: CD4 declined at the negative rate it increased on ART and viral load increased at the positive rate it decreased on ART.

7 INTERHOST EPIDEMIOLOGY

Interhost epidemiological processes simulate the HIV-1 disease transmission in this model. Disease transmission occurs between sexual partners who are active on a given time step. The transmission rate formally provided below is expressed as a rate per partnership per unit time, which was one week in our model. This section will describe how the overall rate as a function of the intrahost epidemiological profile of each member of a partnership, and behavioral features within the dyad.

7.1 Disease-Discordant Dyads

At each time step in the simulation, a list of active dyads is selected from among the nodes in the network. This "edgelist" reflects the work of the network model simulation, wherein partnerships form on the basis of nodal attributes like age and sex, as well as current degree up to the constraint of 3 (see Section 3). Dyads are active at a specific time step if the terminus of that simulated edge is less than or equal to the current time step (right-censored). From this total edgelist, a disease-discordant subset is created by removing dyads in which both members are susceptible or both infected. This leaves those dyads that are serodiscordant on HIV status (one susceptible and one infected partner). This is the set of potential partnerships over which disease may be spread within that time step.

7.2 Per-Act Transmission Probability

Given an active serodiscordant dyad selected from an edgelist, next needed is a per-act transmission probability within that dyad. We use a probability model based on a recent study by Hughes and colleagues on transmissions within serodiscordant couples in chronic-stage infection [31]. Their statistical model relates the per-act transmission probability to the HIV viral load of the infecting partner, as well as the sex, age, and circumcision status of the susceptible partner. Also incorporated in their full model are co-infection with other sexually transmitted infections like Herpes Simplex Virus-2 (HSV-2). Because we did not have individual-level data on STIs, we used the population-level means assumed in the Hughes study.

The core component of the statistical model is the non-linear relationship between the HIV viral load of the infecting partner and the transmission probability given a sexual act. The association is linear in $\log_{10}$ VL, and therefore non-linear with respect to the outcome probability. The probability of transmission from an infected male partner with a $\log_{10}$ VL of 2, 3, 4, 5, 6, for example, to a susceptible female partner aged 35 would be: 0.0001, 0.0002, 0.0005, 0.0015, and 0.0044. The full model is parameterized in the *EpiModelHIV* function `hughes_tp` available on Github (LINK).

As noted, Hughes et al estimated transmission probabilities only for chronic and late-stage infection, and therefore their model does not account for the increased risk of transmission

during acute-stage infection above and beyond that predicted by the peak viremia during that period [20]. Similar to Eaton and Hallet [19] and others modeling primary infection, we added a base multiplier of 5 onto the transmission probability if the infected partner was in the acute stage of infection. This brings the overall transmission probability during this stage in line with earlier predictions from Wawer [32] and Hollingsworth [33], which were better able to quantify the transmission probabilities during early infection.

7.3 Coital Frequency

Following similar models, we used per-partnership act rates based on data from Rakai, Uganda data [32]. The main benefit of this data source is that it includes coital frequency variation by disease-stage, which was not available in MHG. In the absence of stage-specific rates, highly-infective AIDS-stage patients could cause more infections than realistic given their steadily increasing HIV viral loads. Act rates were 0.362 per partnership per day in early and chronic stage infection and reduced to 0.197 during AIDS-stage infection. With these rates, the total number of acts per partnership for that time step were then drawn from a Poisson distribution with the rate parameter set to that respective rate based on the disease stage of the infected partner.

Our initial model included parameters for coital dilution, which is defined as the reduction in per-partnership coital frequency when additional, concurrent partners are added. This has been suggested to have a minimizing effect for concurrency in certain epidemic conditions [34]. Under full dilution, persons have a fixed act budget that does not vary when additional partners are added, which implies that the transmission rate per infected person is invariant to degree (number of ongoing partners). However, our previous work on the MHG study found minimal evidence for coital dilution in Ghana [35], similar to previous reports in South Africa and elsewhere in SSA [36]. Therefore, we did not include dilution in the epidemic model.

7.4 Condom Use

Condom use is typically ignored or conflated into the overall act rate, but we separated it out based on our MHG Study data suggesting that 9% of acts were protected by condoms. To generate a number of protected and unprotected acts at each time step for each serodiscordant dyad, we drew from a Poisson distribution for the total number of acts, and then from a binomial series for each act with a probability equal to the observed condom use. For the effect on transmission risk, we modeled condom use during sex associated with a 78% reduction in the risk of transmission per act, also derived from the Hughes per-act transmission probability model [31].

7.5 Transmission Rate

Given the elements above, the overall transmission rate per partnership per unit time is expressed as:

$$1 - [((1 - \tau * (1 - \omega))^{\alpha_p}) \times ((1 - \tau)^{\alpha_u})]$$

where τ is the base transmission probability, ω is the efficacy of condoms, α_p is the number of acts protected by condoms, and α_u is the number of unprotected acts. Therefore, the overall transmission rate is the 1 minus the product of probability of remaining susceptible due to protected acts and the probability of remaining susceptible due to unprotected acts. These transmission rates were independent across partnerships.

8 INITIAL CONDITIONS AND MODEL CONFIGURATIONS

To initialize the epidemic simulations, it was necessary to set initial conditions of demography and epidemiology from which the natural infectious disease dynamics would flow. For our baseline models, we initialized the prevalence at 1% in order to generate an endemic HIV prevalence and incidence close to the observed value for our data.

Among those initialized as infected, the time of infection was assigned by drawing from a geometric distribution with probability parameter equal to the average treatment-naive lifespan (roughly 12 years). Initial CD4 count and HIV viral load for those infected were based off of the clinical model described above, with the main parameters determining the disease stage linked to age at infection and sex. All infected persons were initialized as undiagnosed and untreated infection, but persons meeting their individualized CD4 threshold for ART were then immediately started upon the simulation.

The time unit for all simulations was one week. The choice of unit in microsimulation models like these is a tradeoff between computational efficiency (longer time units takes less computation given a fixed calendar time period for simulation) and competing risks. The latter is a concern for discrete-time dynamic models, as multiple processes must occur “at” a specific time. The discrete-time representation is an approximation to continuous time, in which individual events occur at distinct times [37]. In our models, a one-week interval for simulation provided the best balance between the bias of competing risks and computational efficiency. Smaller-scale diagnostic simulations at daily time step intervals (not shown) did not yield different substantive outcomes.

In the MHG study population, there was a strong 1.22:1 female-male sex ratio. Typically in mathematical models with sex, a balanced sex ratio is assumed, but we sought to preserve the proportional disparity that was reflective of sex-based differences in demographics, including

death and migration. In our population, high rates of circular migration among men may have led to a selection bias for this sex [14]. Although we did not simulate migration *per se* in the models, it is incorporated within the entry and exit rates along with this imbalanced sex ratio.

9 BURN-IN AND COUNTERFACTUAL SIMULATIONS

The prevailing method for mathematical models for HIV is to model a quasi-historical time series of an epidemic, which typically starts in the early 1980s. Modelers use either fixed behavioral and biological parameters estimated usually more recently than the epidemic onset, or assume they have remained unchanged over that time period [38]. These types of models require, for these and related reasons, parameters for partnership turnover that exceed nearly any population-level empirical data [39]. Other modeling studies use statistical approaches like Bayesian melding [40] to define a plausible range of values for such parameters over that period, and allow them to vary in order to match known epidemiological data on disease prevalence and incidence [41].

Neither approach would be suitable for our models: the fixed-parameter approach because it would be unreasonable to assume that individual behavior related to sexual risk have remained fixed at their current levels since the epidemic's beginning. The Bayesian approach that allows parameter variation over time is undesirable because behavioral and clinical parameters are allowed to vary over multivariate distributions that may not ever match empirical data.

Our approach is to model the *current* state of the epidemic given *current* clinical and behavioral parameters. Therefore, the objective of the simulation is to reach a stable equilibrium point with respect to epidemiology after a burn-in period, and then to analyze results after that stationary distribution has been reached [12]. The burn-in period allows the initialization of the epidemic (e.g., the age distribution and disease prevalence) to not have direct influence on the outcomes under investigation.

For the baseline model representing the epidemic in Ghana, we achieved equilibrium after approximately 50 "years" of simulation. That does not represent the actual time it takes to generate the observed epidemic, since many parameters have changed since those which we observed in more recent empirical data.

Given the stochastic variability in the data, the goal was to generate and maintain a stationary distribution in HIV prevalence and incidence -- that is, an endemic prevalence given the parameters for each model -- across all disease simulations. For this study, we defined the equilibrium point for each parameter set of simulations as maintenance of population-level HIV prevalence within a margin of 0.1% across the final 1000 time steps (~ 20 years).

We used a total of 99 parameter sets. Circumcision prevalence started with the baseline 90% observed in the data, and then down to 10% in increments of 10%. We varied concurrency in a *relative* proportion to the observed, where 100% matched the amount in MHG by sex (Males = 17.8% and females = 2.8%), and 50% was half of that level and 150% was 1.5 times that level. These roughly corresponded to the bivariate levels of circumcision and concurrency that have ever been observed in Sub-Saharan Africa, where the MHG levels are at a relatively high baseline level for both.

For each parameter set, we conducted 100 simulations across 100 years of time. As noted, it took approximately 50 years for the model burn-in period to complete. Therefore, we analyzed data from either the final time step, or an average over the final year of data in each simulation. In total therefore, the sample size for analysis was 9,900 simulations times the number of time steps under consideration.

10 SUPPLEMENTAL FIGURES

Figure S1. Incidence per 100 person-years by relative concurrency prevalence (50% to 150% of baseline values) and absolute circumcision prevalence (10% to 90% of men). Baseline model is 100% relative concurrency and 90% circumcision. Brightest red indicates highest incidence and dark grey lowest incidence.

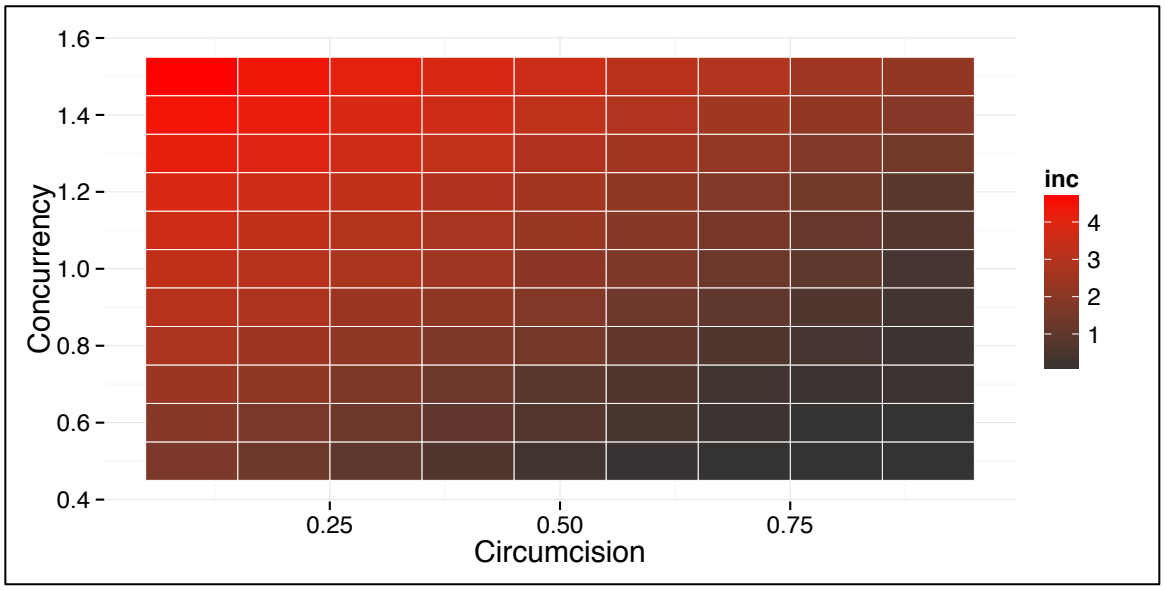


Figure S2. Probability of epidemic extinction by relative concurrency prevalence (50% to 150% of baseline values) and absolute circumcision prevalence (10% to 90% of men). For the model in which relative concurrency was 50% and absolute circumcision was 90%, 28.6% of epidemics went extinct; purple cells adjacent show that 7.4% of epidemics with those values were extinct.

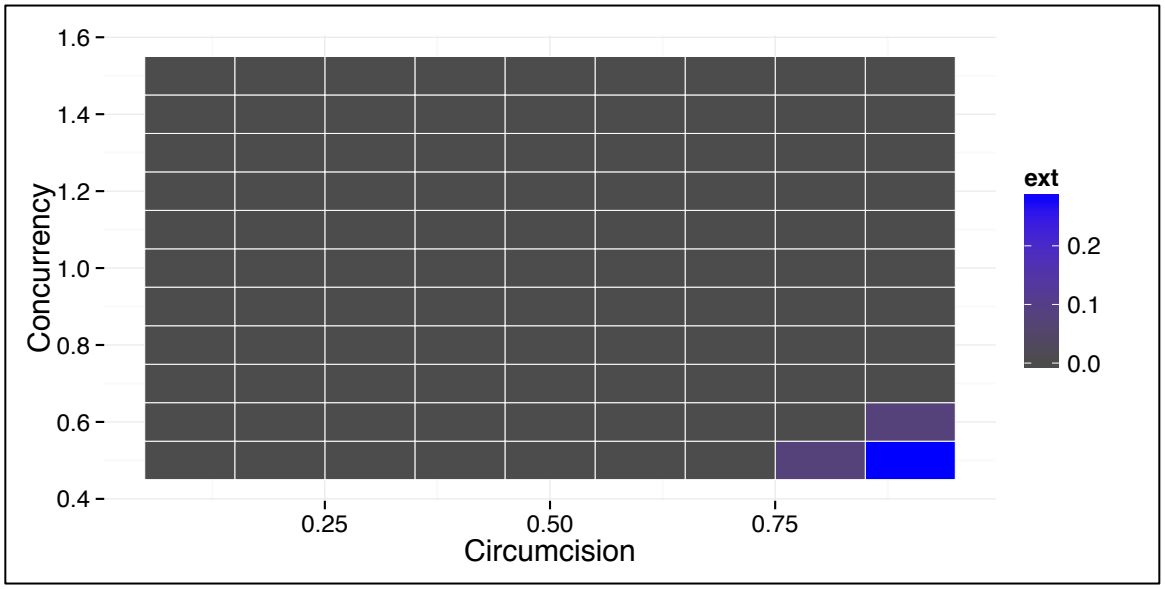


Figure S3. Incidence rates per 100 person-years over burn-in and analysis time periods. Rates show baseline concurrency levels with counterfactuals for absolute prevalence of male circumcision (90% to 10%). The burn-in period lasts for the first 2,600 time steps (50 years), and the analysis period lasts for the final 1,000 time steps (~20 years) at which point incidence has stabilized.

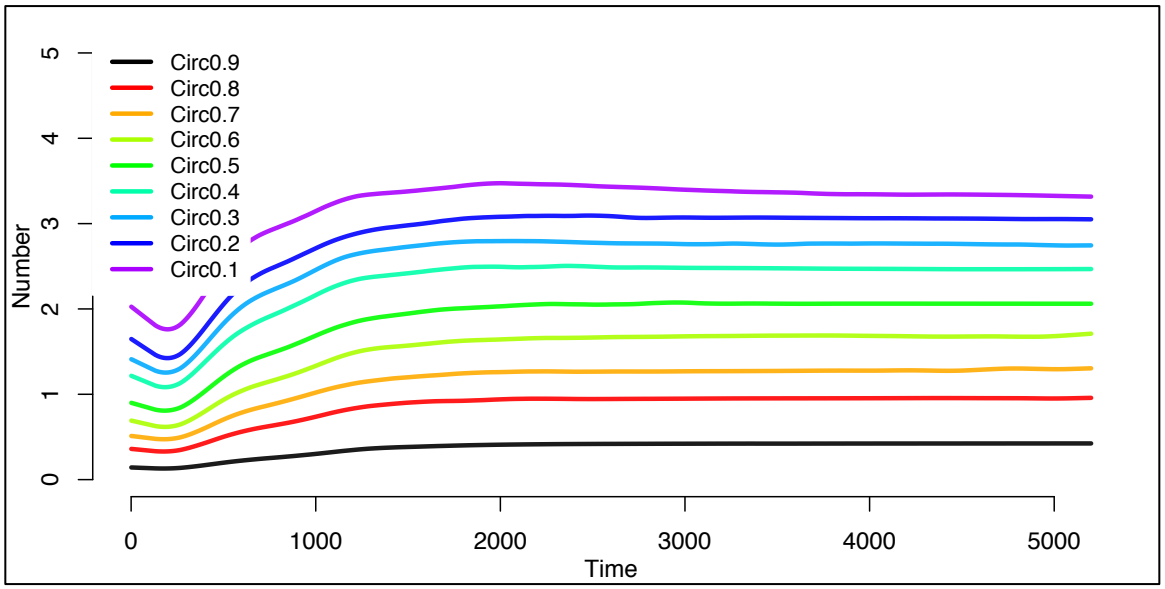


Figure S4. Endemic HIV prevalence by relative concurrency level (50% to 150%) with 50% circumcision prevalence among men (left panel) and 10% circumcision among men (right panel). Boxplots show the distribution of outcomes for the final 100 time steps for each simulation set (100 simulations each) for 11 counterfactual models. The red line is a lowess smoothed curve over the means of across concurrency levels.

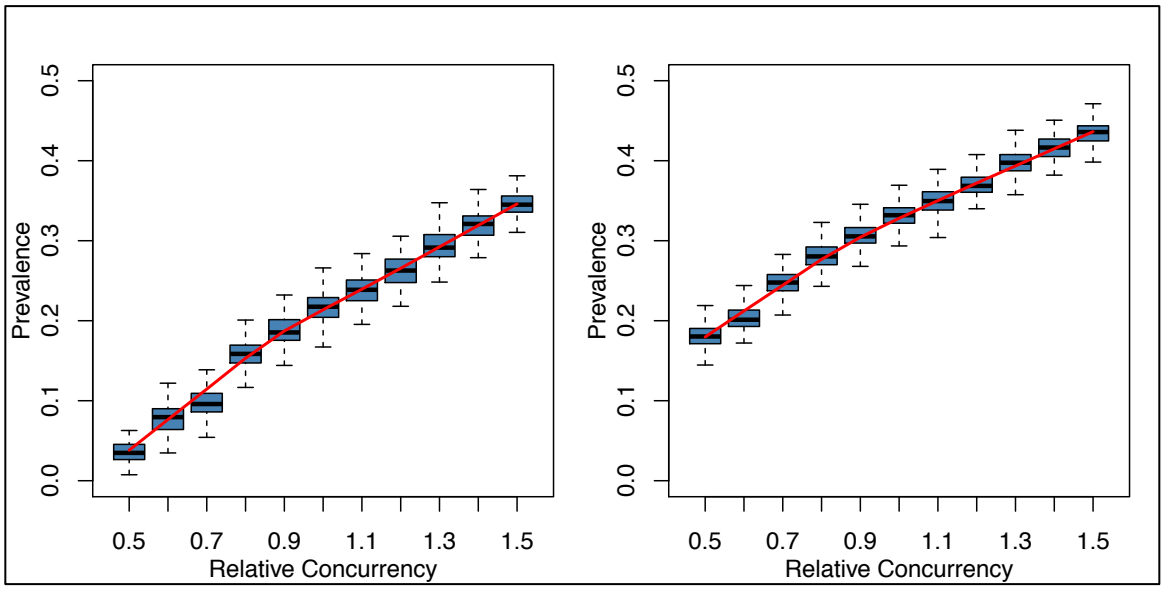
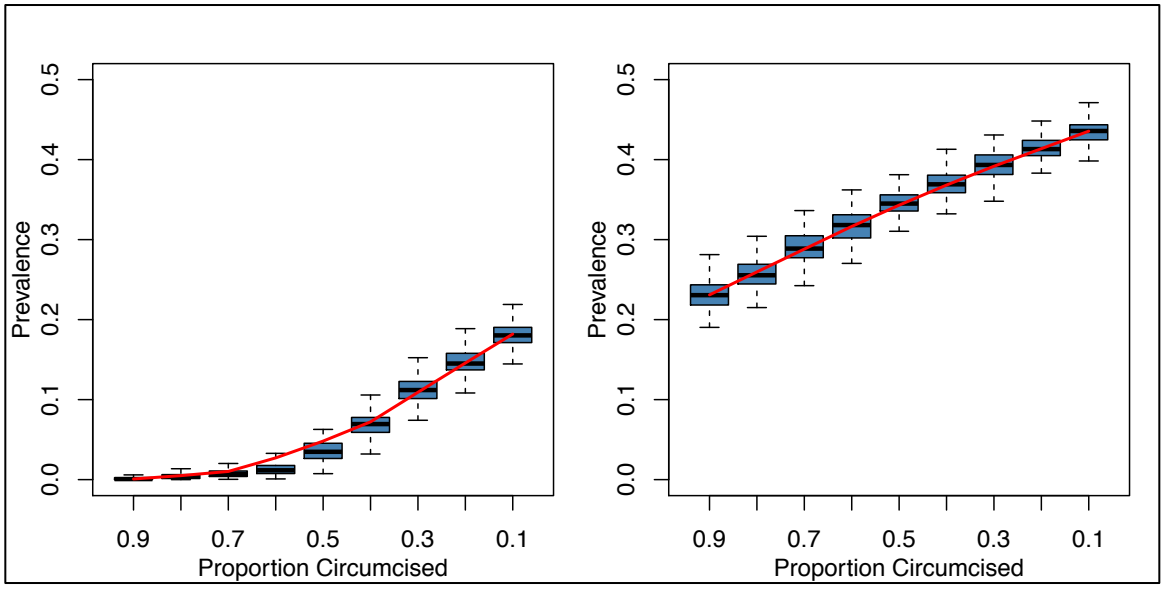


Figure S5. Endemic HIV prevalence by absolute circumcision level (90% to 10%) with 50% relative concurrency prevalence (left panel) and 150% relative concurrency prevalence (right panel). Boxplots show the distribution of outcomes for the final 100 time steps for each simulation set (100 simulations each) for 11 counterfactual models. The red line is a lowess smoothed curve over the means of across concurrency levels.



Chapter 3: Interference for Pre-Exposure Prophylaxis HIV-1 Prevention Efficacy in Dynamic Contact Networks

Interference for Pre-Exposure Prophylaxis HIV-1 Prevention Efficacy in Dynamic Contact Networks

Samuel M. Jenness et al.

March 12, 2015

ABSTRACT

Background Randomized controlled trials (RCTs) of prevention agents for infectious disease may be subject to interference if the probability of infection for any trial participant depends not only on that person's treatment status that of others. Interference can bias the direct effects observable in RCTs in study settings with high sexual network connectivity. This may partially explain the variability in RCT results for oral pre-exposure prophylaxis (PrEP) among heterosexuals in Sub-Saharan Africa.

Methods With stochastic mathematical network models for HIV-1 transmission dynamics, we simulated a high-incidence heterosexual HIV epidemic in which an RCT of PrEP for high-risk women occurred, parameterized with behavioral data from recent trials. Time to seroconversion or censoring across simulations was analyzed using survival methods, with counterfactual scenarios varying PrEP adherence, drug efficacy, RCT coverage level, and network connectivity.

Results Model scenarios with PrEP adherence at 25% replicated the null findings from recent RCTs. In the baseline full adherence scenario, interference biased the observable treatment efficacy hazard ratio by 2%, with unobservable indirect effects responsible for 10% of the total effects of PrEP in the community. The greatest impact was achieved when RCT coverage was highest, resulting in a 4% bias to the hazard ratio and indirect effects comprising 25% of total effects. Indirect effects were minimized when network connectivity, and therefore background disease incidence, was lower.

Conclusions Indirect effects comprised a non-trivial proportion of the total effects of PrEP to prevent HIV infection under certain RCT conditions, but the absolute bias to the observable hazard ratio was minimal. Currently designed PrEP RCTs are therefore robust to the interference challenge, and poor adherence remains the primary factor driving variability in trial results to date.

1 INTRODUCTION

Interventions for infectious disease (ID) prevention can result in nonlinear effects on disease incidence because of the possibility of protection via indirect effects [1]. In contrast to most non-infectious diseases where individual-level risk is modified only by the presence of an individual-

level exposure, ID interventions can confer protection to the unexposed [2]. A broad goal of dynamic mathematical modeling of epidemics is to quantify the population-level implications of these indirect effects, often under the banner of "herd immunity" for vaccine-preventable diseases [3].

Indirect effects also impact epidemiologic studies establishing their efficacy. What is often desired from a randomized controlled trial (RCT) is a comparison of disease incidence among individuals in the presence or absence of an intervention, holding all else constant through randomization. But what is observable in even the best executed ID trial may be a mixture of these direct effects along with indirect effects [4]. This happens when an intervention confers prevention benefits to those in the control arm; it is a function of how persons in each arm circulate infection to each other within the larger community from which both were recruited [5]. Evaluating the same intervention (with unknown efficacy) in two study populations may yield divergent results. Within the counterfactual literature in epidemiology, this is called interference, or a violation of the "stable unit treatment value assumption" (SUTVA) [6]. Methods to quantify interference and adjust study results for it are needed when solutions aimed at the study design, such as community-level randomization, are impossible or infeasible [7].

How interference affects observable treatment effects within a trial environment depends on the contact patterns between persons in the study [8]. For analytic tractability, the interference literature to date has primarily focused on patterns that are relatively basic (persons within households) and fixed over the study follow-up time [9], motivated by vaccination campaigns consistent with these assumptions [10]. There has been little attention paid to contact structures with dynamic and complex mixing patterns, partially because deriving analytic solutions may be impossible [11]. Sexual partnerships are one such contact structure, with RCTs of interventions for the prevention of HIV and sexually transmitted infections a needed area for interference research.

Oral pre-exposure prophylaxis has generally been shown to be an effective agent for the prevention of HIV-1 infection in RCTs among several high-risk populations [12-14]. Within heterosexuals in Sub-Saharan Africa, however, there have been discrepant results. The Partners PrEP study of serodiscordant couples found an intent-to-treat effect of 75% [12], while the Fem-PrEP and VOICE trials observed no effect [15, 16]. One notable difference between the trials is drug adherence, with the positive trials maintaining greater than 70% adherence and the null trials less than 40%. But there are several other differences related to the study design that could explain the diluted efficacy in the null trials. Fem-PrEP was conducted among young women with high levels of partner change and concurrency (multiple overlapping partnerships [17]), whereas Partners PrEP was among stable serodiscordant couples in long-term

partnerships with lower concurrency. The baseline incidence rates in Fem-PrEP and those in the control arm of that trial were over twice that of Partners PrEP.

In this study, we investigate how interference impacts the observable results of an RCT designed like Fem-PrEP. Our hypothesis is that the structural elements of the sexual network in a high-incidence setting will recirculate disease rapidly enough to confer non-trivial indirect effects that reduce observable differences in incidence between treatment and control arms, diluting the hazard ratio towards the null. Whereas prior research on interference has focused on highly transmissible infections like pneumococcus [9], ours will be one of the first to quantify this for a sexually transmitted infection. Given the complexity of the dynamic contact network, we use a stochastic mathematical network model of HIV-1 transmission dynamics in which elements of the RCT, including population coverage, treatment efficacy, and drug adherence are varied along with network connectivity. This allows partitioning of the total effects, invariant across study milieu, into direct and indirect effects to estimate the impact of interference.

2 METHODS

This study uses a stochastic network model of HIV-1 transmission dynamics to simulate an HIV-1 epidemic among heterosexuals in Sub-Saharan Africa (SSA). The model includes background disease transmission in the population before the initiation of an RCT, which occurs at a set time point for a defined period of follow-up. The simulated network structure follows specified patterns of sexual partnership formation and dissolution over time, with model parameters derived from empirical data. The full methodological framework for this study, including network model estimation, as well as epidemic model parameterization, simulation, and data analysis, are provided in the Supplementary Appendix **[LINK]**.

Network Model Form

The dynamic sexual contact network generating this epidemic is based in the statistical framework of temporal exponential random graph models (ERGMs) [18]. ERGMs provide a mechanism for the simulation of complete network over time given a set of network statistics hypothesized to govern partnership formation and dissolution. Targets for these network statistics are estimable from egocentric network data [19].

Parameters for the network model are largely drawn from secondary data from the Fem-PrEP trial [15]. We assumed that the network structure for persons within the RCT was the same as those from the underlying population of sexually active persons, the implications of which are discussed in the study limitations. This was a heterosexual transmission model, and therefore men only partnered with women in the population, and vice versa. Mean degree, or the average number of ongoing partnerships per person at any time, was set to one based on the proportion

of study participants reporting a primary partnership. Table 1 lists this and other parameter values for the network model. At baseline, 26% of women in Fem-PrEP had a secondary partner, so we set the level of concurrency to this level. This statistic was entered into the network model as the product of the number of women in the population and this proportion. Finally, we modeled sex-structured age homophily. Typical of heterosexual mixing [20], persons selected partnerships similar in age (age homophily) with a variance in the absolute difference in ages based on our prior models of HIV-1 transmission in SSA heterosexuals.

The model also required a dissolution component that is a function of the mean partnership duration. This mean may be estimated using Kaplan-Meier or related survival methods that account for the right-censoring of partnerships ongoing at the end of the study period [21]. For our current purposes, the average duration was treated as a parameter to calibrate the HIV incidence in our simulations to observed levels in the trial. Incidence is negatively correlated with mean partnership duration since infection becomes locked within longer-duration partnerships [22], so we calibrated the coefficient controlling mean duration downward (mean = 380 days) until the epidemic reached an equilibrium HIV incidence rate of 5 per 100 person-years in the subset of women in the population with an age range matching the eligibility criteria of Fem-PrEP [15].

HIV Progression and Transmission

The focus of this model is a simulated RCT of a subpopulation of initially disease-free women, with follow-up to determine their rate of infection through heterosexual activity. It is therefore necessary to model the dynamics of HIV transmission in the full population of men and women for infections within the RCT subpopulation to occur. Our model simulates a set of mechanics for HIV progression given the natural course of infection and anti-retroviral therapy (ART) treatment profiles similar to prior approaches [23]. Persons progress through HIV stages with evolving CD4 and HIV viral loads, which are dependent over time. Infected persons initiate treatment based a randomly assigned CD4 count for ART initiation from our prior models [24]. Upon ART, persons were partitioned into full and partial adherence groups such that the average levels of HIV viral suppression matched broad empirical estimates for SSA [25]. ART adherence was associated with an increase in CD4 to pre-infection levels and a reduction in VL to suppression [23, 24, 26].

HIV was transmitted over active partnership dyads given the network model structure at each time step. The per-partnership transmission rate started with the Hughes [27] statistical model for the per-act transmission probability, which linked it to the HIV viral load of the infected partner, condom use, as well as the sex, age, and circumcision status of the susceptible partner. The final per-partnership transmission rate was a function of the per-act transmission probability and the number of acts per unit time.

Baseline Conditions

Baseline incidence in the full population was established through a set of burn-in simulations. Initial values for the proportion infected, along with HIV-related attributes such as time since infection, were set at the initial time step of the model. The starting network size was 25,000 persons. A total of 100 burn-in simulations were then used to establish of demographic, epidemiologic, and behavioral attributes at an equilibrium state. The result of this series was a set of simulations with these attributes, along with HIV incidence, in this stationary distribution. From this set, we selected the median simulation with respect to disease incidence for the RCT simulations. A total of 250 simulations were run for each counterfactual scenario described below.

Randomized Control Trial

The RCT simulation started by recruiting age-eligible women, and followed them to monitor seroconversions to establish rates of HIV incidence by study arm. During a recruitment window of 2 years, eligible women were randomly recruited from the population at a rate set to achieve the desired sample size within the window. Upon study entry, women were randomly assigned to treatment or control arms and then tested for HIV infection at monthly intervals. The primary study end point was documented seroconversion. Subjects could also exit the trial due to loss to follow-up and mortality. Censoring was also possible if women reached a maximum follow-up time of 36 months. This observation window length was based on the recent VOICE trial [16] instead of Fem-PrEP because the latter trial had a very truncated follow-up window due to early stoppage for futility. Individual-level data for time on study and endpoints was tracked for each simulation.

RCT participants on the treatment arm were randomly assigned an adherence type corresponding to full or partial use of PrEP, similar to our ART adherence model. Full adherers stayed on PrEP for the entirety of their study follow-up, while partial adherers cycled on and off drug as a Markov process at each time step. Adherence levels were calibrated such that the observable adherence to study medication at diagnostic visits matched specified levels, including the low levels observed in Fem-PrEP. PrEP use conferred a reduction in the per-act risk of infection. Partial adherers were protected during those intervals of use, and unprotected in intervals of non-use. Baseline levels for this efficacy measure were drawn from a review of per-protocol effects within PrEP RCTs [28], and efficacy was varied in counterfactual scenarios.

Data Analysis

Simulated RCT data were analyzed using Kaplan-Meier methods to estimate the non-parametric survival distribution, along with calculations of within-group incidence rates. The stochastic

variability in the incidence rates and related hazard ratios comparing treated to placebo arms was quantified through the variability in outcomes across the 250 simulations per scenario.

A community control trial was also simulated by recruiting from the same set of women, followed for the same time, but with treatment efficacy of PrEP set to null. This allowed for estimation of background incidence among a fully unexposed population with no indirect effects. Following the methods of Halloran and colleagues [2, 4], the direct effects were estimated as the complement to the hazard ratio (1 – HR) comparing incidence in the treatment and placebo arms in the trial population, while the indirect effects was the similar expression of the comparison between incidence of the community control against the placebo arm in the trial population. The total effects of PrEP were therefore calculated as one minus the hazard ratio comparing community control against the treatment arm in the trial population.

3 RESULTS

Across all simulations on this 25,000-person network, the community control group simulation with no exposure to PrEP maintained a mean incidence of 5.01 per 100 person-years in the sex and age defined eligibility group. Figure 1 shows the Kaplan-Meier survival curves in the treatment and placebo arms of the RCT population across all simulations under two scenarios. The left panel shows the incidence in the baseline model in which there was perfect adherence to PrEP that had a per-act treatment efficacy of 80%. The right panel shows the same survival distribution in a counterfactual with the adherence levels similar to the Fem-PrEP trial (25%). In this low adherence scenario, the survival probability was lower and therefore the incidence higher overall in the placebo arm compared to the treatment arm, but the large overlap of the distribution of individual curves indicates the improbability of observing a significant treatment effect in a single trial.

Table 2 provides the main efficacy results, partitioning the incidence rates in the treatment and control arms into their direct and indirect effects through comparison to the community control. The baseline scenario for counterfactual comparisons was 100% adherence in which 20% of the eligible population was recruited into the trial, with a treatment efficacy of 80%, and with the levels of network connectivity in sexual partnerships observed in the Fem-PrEP study. In the baseline model, mean incidence in the treatment arm was 1.02 per 100 person-years, compared to 4.61 in the control arm and 5.01 in the community control group.

The PrEP intervention was therefore associated with a direct effect of 78%, comparing the two arms in the study population, and an indirect effect of 8% comparing the control arm to the community control. The total effect of the intervention matched the target efficacy of PrEP at 80%, as expected. The direct effect therefore was a biased estimator of the total effect by 2% in absolute terms because the intervention had resulted in this 8% relative decline in incidence in

the control arm with the RCT community. Overall, this indirect effect contributed 10% of the total effect of PrEP in the RCT population, whereas the direct effect contributed the remainder.

The relative contribution of the indirect effect grew when adherence to the study drug decreased, as shown by comparing the indirect ratios across these adherence scenarios. This was driven by the lower levels of adherence that grew the incidence rates in the treatment arms at a greater level than they did in the control arms. While the Table lists the mean results, the left panel of Figure 2 shows the stochastic variation incidence rates across all simulations. In the baseline model, the observable direct effect was a biased estimator of the total effect by 2%: under no interference the direct effect would be 80%, but the presence of indirect effects reduced it to 78%. In the low-adherence scenario, this bias was 3% (15% observed efficacy versus 18% expected). The modeled treatment efficacy of the drug also changed the magnitude of the indirect contribution to the total effect. Higher efficacy treatment was associated with indirect effects of 11% of the total assuming PrEP were 90% effective, whereas it contributed 2% if PrEP were 50% effective. More effective interventions had a greater impact on the incidence in the treatment group relative to the incidence in the control group, widening the gap in incidence between observable treatment arms.

Coverage level of the RCT, defined as the proportion of eligible women in the population recruited into the trial, also contributed to indirect effects. If 10% of the population were recruited into the trial, the indirect fraction was 6%, whereas if 50% of the population were recruited it was 25%. Recruitment of a higher fraction of the population resulted in a larger reduction in the control arm of the RCT. This resulted from the prevention of more infections within the full population in which the RCT occurred, leading to a decline in risk for study participants in the control arm. Increasing the recruitment coverage also had a minor influence on the incidence rates within the treatment arm, since at highest coverage levels the infection risk of those on treatment was reduced as a function of the lower background risk of acquiring infection even with PrEP use (this would not occur if PrEP were 100% effective).

Indirect effects were also greater in high-incidence settings generated by higher levels of network connectivity, compared to low-incidence settings with lower network connectivity. When the level of concurrency in the population was reduced by half, the incidence rate in the control community fell to less than half of what it was in the baseline model (2.11 per 100 person-years). With a similar baseline incidence in the RCT community, PrEP continued to have both direct and indirect effects on incidence: a 78% reduction in incidence comparing the treatment and control arms and a 2% reduction in incidence comparing the control arm to the control community. This lower incidence setting therefore resulted in a decreased relative contribution of indirect effects, and a smaller bias to the observable estimate.

4 DISCUSSION

In this study, we used stochastic network models for HIV-1 transmission dynamics to investigate how interference under difference scenarios could bias the observable direct effects of an oral PrEP trial for HIV prevention. The relative contribution of indirect effects within the RCT depended on adherence levels in the trial, the coverage of the trial within the broader population, the treatment efficacy of the study agent, and the levels of network connectivity generating the baseline HIV incidence prior to RCT enrollment.

Indirect effects contributed a maximum of 25% of the total effects of PrEP across scenarios, yet this only resulted in a bias to the observable hazard ratio by 4% in absolute terms. Therefore, we conclude that community-level indirect effects of treatment are unlikely to explain much of the variability in PrEP trial results to date. Our model confirms that varying levels of adherence is the strongest predictor of bias to the intent-to-treat estimator, with the magnitude of bias negatively correlated with the adherence level, similar to findings using both routine per-protocol analyses [28] and more rigorous principal stratification methods [29].

Although our study suggests that the currently implemented PrEP trials would not be subject to significant bias due to interference, these results have implications for the design of future HIV prevention trials. For highly effective prevention tools like PrEP and ART, a less biased estimate of the treatment effect will be achieved when recruiting a smaller fraction of the population with lower baseline incidence. This presents a tradeoff to cost-effectiveness, since this would require observing more person-time spread over a larger population (e.g., multi-site studies) to quantify an unbiased effect. If indirect effects are large enough under some scenarios, a community RCT like that of the ongoing PopART study may be necessary [7].

Compared to previous research on interference with applications to more highly contagious respiratory and vector-borne diseases [9], our study suggests that HIV prevention trials are less susceptible to interference-related biases. Even with highly connected sexual networks with relatively rapid partnership turnover, HIV simply does not circulate within the RCT community quickly enough during the trial observation window to impact the study results greatly. In comparison, a set of trials for an oral cholera vaccine found that the incidence rate varied almost 7-fold for the placebo arm and over 2-fold for the treatment arm when comparing communities in which vaccine coverage was greater than 50% to communities where it was less than 28% [30]. Compared to HIV in a partially ART-adherent population with an R_0 approaching 1, the higher base infectivity of cholera generates higher levels of interference for a prevention study [31]. Although our model did not address this, observation time is likely also quite important. A recent model of combination packages for HIV prevention suggested that the population-level effects of the intervention are not realized until long after a typical observation window of an RCT [32]. Therefore, RCTs with greater follow-up time may be subject to greater interference.

Methodologically, our approach of a stochastic mathematical network model to address these questions suggests a promising alternative to the quantification of direct and indirect effects through community RCTs, which may be infeasible or unethical [33]. The ethical equipoise for PrEP has been lost, and a control condition would not be warranted: current demonstration projects for PrEP use simulation modeling to generate control counterfactuals to compare against active PrEP scale-up arms [34]. Compared to deterministic compartmental models, the network-modeling framework is particularly useful for investigating these phenomena because epidemic scenarios may be analyzed by starting with empirically grounded biological and behavioral parameters, drawing on the epidemic acceleration possible in dynamic, highly connected networks.

Limitations

One limitation of our model is an assumption that the sexual risk behavior of the RCT enrolled subgroup is the same as the broader population. In other words, there is no selection bias in sampling from the full population generating the HIV epidemic. That assumption is currently required primarily due to the lack of population-level data on the risk and clinical profile of the full population, but also due to complexities with modeling heterogeneous network configurations (e.g., multiple partnership types) within one epidemic model. We also do not model behavior change resulting from RCT participation. The Partners PrEP study, for example, observed a decline in those behaviors in both treatment and control arms due to extension behavioral counseling at study visits [12]. Our future work on this subject will explore the impact of both selection bias, behavioral risk reduction, as well as network and risk heterogeneity, on the levels of interference within a trial setting. It is likely that the indirect effects are further minimized if the broader population is at lower disease risk, and if risk reduction occurs within study participants, due to delayed recirculation of disease.

Another limitation is that we were unable to disentangle the changes to network connectivity from changes to the baseline HIV incidence rates in the population. More highly connected networks (i.e., with more concurrency) generate higher disease incidence because the virus moves through them more quickly than networks in which there is only serial monogamy [17, 35]. Varying the prevalence of concurrency but holding incidence constant would require changing other components of the dynamic network structure, such as the average partnership duration, or changing a biological or behavioral risk factor, such as the rates of coital acts. Therefore, counterfactuals on network structure are intertwined with background incidence rates at present.

Conclusions

This study provides reassurance that the RCTs of PrEP are not substantially biased by interference, which occurs when the probability of infection for any trial participant depends on the treatment status not just of that individual but of others in the study. Interference generates indirect effects that can dilute the observable direct effects within a trial, but we found only small biases (<4% to the hazard ratio) under the extreme scenarios of RCT coverage. The problem of poor adherence to PrEP continues to plague both clinical trials and implementation research in many high-risk populations [16]. Network modeling of these mechanics provides a cost-effective, efficient framework to establish the potential for biases to treatment effect estimation within HIV prevention RCTs, but only if grounded in solid empirical data on behavioral and biological data.

5 TABLES

Table 1. Main Model Input Parameters

Parameter	Value
RCT Configuration	
Age Eligibility	18-35
Recruitment Rate (weekly)	0.039
Recruitment Period	24 months
Diagnostic Test Interval	Monthly
Loss to Follow-up Rate (weekly)	0.00012
Maximum Follow-Up Time	36 months
Treatment Arm Fraction	0.5
Coital Frequency	
Monthly acts	14.8
Monthly unprotected acts	7.6
Network Model	
Mean degree	1.0
Proportion concurrent (high)	26%
Proportion concurrent (low)	13%
Mean partnership duration	380 days

Table 2. Incidence Rates per 100 Person-Years and Partition of Total Effects into Direct and Indirect Effects, Across Counterfactual Scenarios

Conditions	Incidence per 100 PY			Effects Partitions			
	Treatment	*Control*	*Comm. Control*	*Direct*	*Indirect*	*Total*	Indirect / Total
Adherence							
100% (baseline)	1.02	4.61	5.01	0.78	0.08	0.80	0.10
75%	2.42	4.68	5.01	0.48	0.07	0.52	0.13
50%	3.57	4.75	5.01	0.25	0.05	0.29	0.18
25% (obs.)	4.11	4.83	5.01	0.15	0.04	0.18	0.20
RCT Coverage							
10%	1.03	4.79	5.01	0.78	0.04	0.79	0.06
20% (baseline)	1.02	4.61	5.01	0.78	0.08	0.80	0.10
30%	0.99	4.44	5.01	0.78	0.11	0.80	0.14
40%	0.99	4.20	5.01	0.76	0.16	0.80	0.20
50%	0.98	4.01	5.01	0.76	0.20	0.80	0.25
Treatment Efficacy							
0.9	0.53	4.50	5.01	0.88	0.10	0.89	0.11
0.8 (baseline)	1.02	4.61	5.01	0.78	0.08	0.80	0.10
0.7	1.52	4.77	5.01	0.68	0.05	0.70	0.07
0.6	2.01	4.87	5.01	0.59	0.03	0.60	0.05
0.5	2.50	4.96	5.01	0.50	0.01	0.50	0.02
Network Connectivity							
High (baseline)	1.02	4.61	5.01	0.78	0.08	0.80	0.10
Low	0.45	2.06	2.11	0.78	0.02	0.79	0.03

6 FIGURES

Figure 1. Survival distributions for baseline simulation with full adherence (left panel) and counterfactual scenario with adherence matching the Fem-PrEP trial (right panel). Thin lines represent 250 simulations of the trial and thick lines are the mean survival distributions across trials. The means in disease incidence in the low adherence trial diverge, but the probability of observing a difference in any one trial is low given the overlap between the simulation lines.

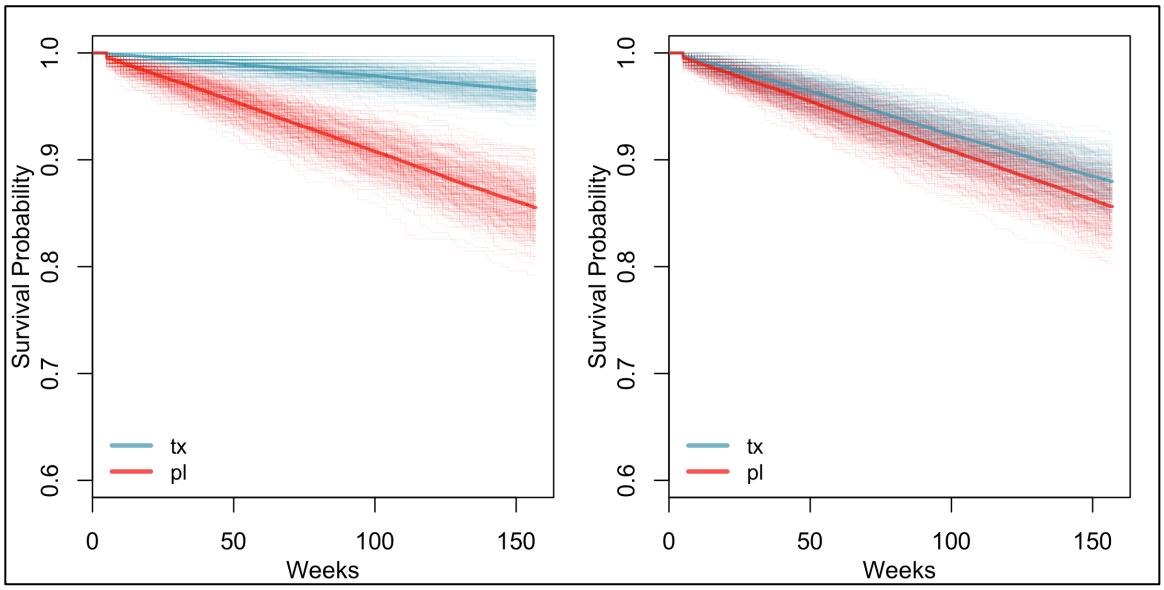
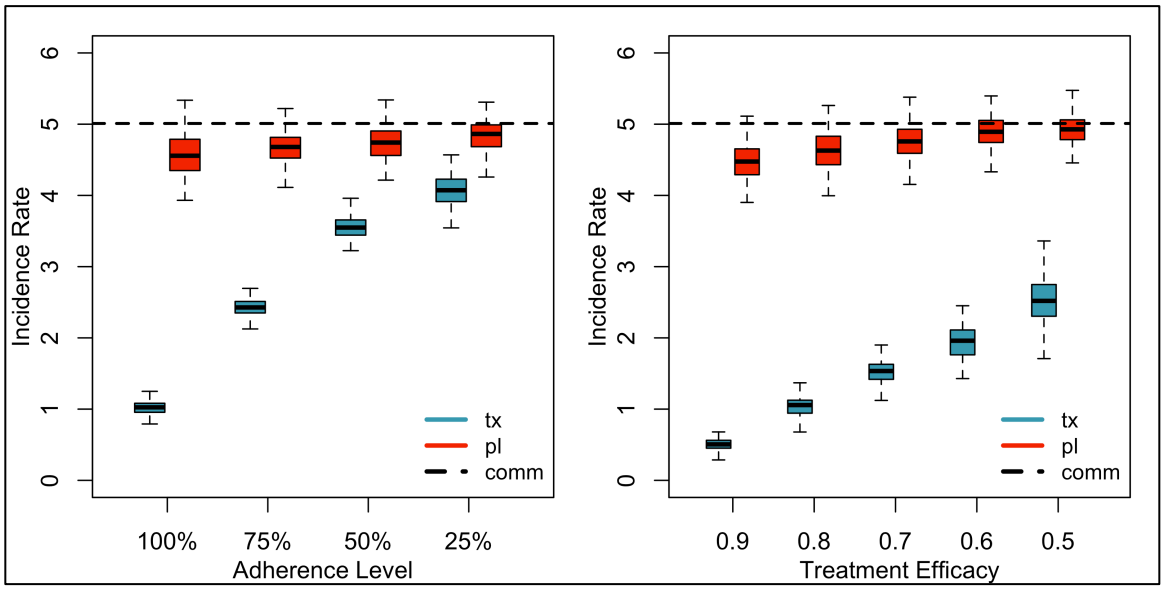


Figure 2. Incidence rates in the treatment arm, placebo arm, and community control in counterfactual scenarios varying adherence level to PrEP and treatment efficacy of the agent. The incidence rate in the community control is invariant as there was no drug available. Higher adherence is associated with a decreased disease incidence in both the treatment and control arms, while treatment efficacy is negatively correlated with incidence in both arms.



Interference for Pre-Exposure Prophylaxis HIV-1 Prevention Efficacy in Dynamic Contact Networks

Supplementary Appendix

Samuel M. Jenness et al.
March 12, 2015

1 INTRODUCTION

This supplementary appendix describes the mathematical model structure of the accompanying paper in more detail. The models presented in this paper are individual-level microsimulation models in which uniquely identifiable sexual partnership dyads are simulated and tracked over time. This partnership structure is made possible through the use of temporal exponential-family random graph models (ERGMs), which are described in detail below.

Along with this dynamic network simulation, the larger epidemic model includes demography (entries, exits, and aging), interhost epidemiology (disease transmission), intrahost epidemiology (disease progression), and clinical epidemiology (disease diagnosis, treatment initiation, and adherence). The main addition to our prior models is pre-exposure prophylaxis as an agent for prevention of HIV infection, which for this study is integrated this into a randomized control trial (RCT) module. The RCT module simulates HIV transmission dynamics within and outside of the trial environment, so track changes to incidence rates at the individual, study, and community levels attributable to the agent. Individual attributes related to these processes are stored and updated in discrete time over the series of time steps for each simulation.

Computationally, this model was programmed in R and C++ using *EpiModel* (www.epimodel.org), which is software developed by the authors for the purpose of simulating complex network-based mathematical models of infectious diseases, primarily HIV and other sexually transmitted infections. *EpiModel* is part of and based on *Statnet*, which is a suite of software in R that provides tools for the representation, visualization, and statistical analysis of network data [1].

This current model extends the general *EpiModel* disease platform to incorporate HIV-specific elements not part of the core software tools. It will be released as a stand-alone R package, *EpiModelHIV*, on both Github (LINK) and CRAN (LINK).

2 CONTACT NETWORK MODEL

At the center of this model is the dynamic network structure, which concerns the HIV-related sexual contacts that individuals make over time. A dynamic model for the formation and dissolution of sexual partnerships based on the empirical data above was formed with the

statistical methods of exponential-family random graph models (ERGMs) [2]. ERGMs provide a foundation for a statistically principled simulation of local and global network structure given a set of target statistics from empirical data. In this section, we describe the statistical and mathematical framework of this approach, and then discuss the unique parameterization of this model.

2.1 Mathematical Framework

Temporal ERGMs consist of a formation model and a dissolution model. The formation model governs how two actors not previously connected within the network pair, and the dissolution model influences how two currently paired actors dissolve their dyad. In the formation model, the probability of observing a set of *Y* relations among *n* actors given a set of attributes is expressed as:

$$\mathrm{P}(\mathrm{Y} = \mathrm{y}|\mathrm{n}, \mathrm{X}) = \exp\left(\frac{\theta' \mathrm{z}(\mathrm{y})}{\kappa(\theta, \mathrm{n}, \mathrm{X})}\right)$$

where $\mathrm{z}(\mathrm{y})$ are a set of network statistics, and θ are coefficients to be estimated from the model. The denominator of the equation is a normalizing constant, which given the potential number of network configurations of a sufficient size, is not estimable. Therefore, simulation-based estimation methods using Markov-Chain Monte-Carlo (MCMC) sampling are used to obtain maximum likelihood estimations for θ given the set of network statistics, z [3].

Network statistics may be of arbitrary complexity, and fall into two main classes: dyadic independent and dyadic dependent statistics [2]. Independent statistics are those in which the probability of dyad formation between any two nodes, i and j, does not depend on the existence of ties from i and j to other nodes (i.e., persons) in the network. Examples include the propensity for mixing by sex (i.e., heterosexual mixing) or age-based homophily (i.e., choosing a closely aged partner). In contrast, dependent statistics are those in which the probability of the tie between i and j depends on the presence of any ties from those two nodes to other nodes in the network. A common example is degree, which is the number of active ongoing ties for each node. The probability of forming a tie with a new potential partner may be hypothesized to depend on ties already in existence.

Dynamic network models also include a dissolution component that predicts the network statistics associated with ties dissolving conditional on their existence [4]. Within each time step, the formation and dissolution models are said to be separable: for any node the process of tie formation is independent of tie dissolution within a time step, but dependent over time. The implications of this assumption are that the size of the time step must be minimized to reduce the effects of the within-step independence.

The dissolution model may be estimated a number of ways. In this current study, we use the Edges Dissolution approximation method of Carnegie et al [5], in which a static ERGM was fit and the formation coefficients statistically adjusted to account for edge dissolution. This method was chosen for computational tractability compared to fitting full temporal ERGMs. The two methods are implemented in the EpiModel software.

With this method, a complete network need not be observed for the purposes of dynamic simulation. We used egocentrically based targets to query localized network measures (e.g., attribute-based homophily and degree). The MCMC-based simulation used for network model estimation is also used for the simulation of complete networks with features stochastically varying around target statistics from the data. The specific parameterization is provided below, but the egocentric framework is discussed in more detail elsewhere [6].

Finally, for the dynamic simulations two adjustment methods were used to account for the changes to the population composition over time. The first adjustment was to preserve mean degree (i.e., the average number of partners per person at any time) given fluctuations in population size with births and deaths. The offset method of Krivitsky and colleagues [6] holds this mean degree constant, a reasonable assumption for HIV/STI epidemic models. This adjustment also allowed the generation of a network sufficiently large enough to represent the internal dynamics of the RCT setting within the population. We started all simulations with a network size of 25,000 nodes, which remained stable over time.

The second coefficient adjustment method was to account for the presence of deaths or exits from the network as an exogenous competing risk for partnership dissolution, on top of the endogenous rates of partnership dissolution observed in the data. We used an adjustment of the dissolution coefficients whereby the probability of tie persistence (the complement to tie dissolution) was adjusted upward to account for the competing risk from mortality.

2.2 Parameterization

The contact network model was parameterized based on secondary data available from the Fem-PrEP study [7]. This section provides information on the structure of the formation and dissolution models, how the data were estimated, and methods for model calibration.

Formation Model. Given the aims of this study, our network models were formulated to capture the sex and age-based mixing structure and network connectivity in a heterosexual population similar to those in the Fem-PrEP trial (the details of which are described in Section 8 below). The network statistics were generated through egocentric inference from the empirical data. The formation model included the following elements:

1. **Mean degree.** The overall number of dyads observed in the cross-sectional data was modeled with an `edges` term in the ERGM. This term is commonly used like an intercept in

a generalized linear model, so that the remaining coefficients are interpreted in its base value. The number of edges in the model was based on a calculation of cross-sectional mean degree at baseline in Fem-PrEP [7]. As noted, the coefficient was adjusted to account for differences in the observed study sample and the simulated network size, preserving the mean degree. In the data, the observed mean degree was 1 overall, corresponding to an expected 12,500 edges in a network of size 25,000.

2. **Heterosexual mixing.** This model is designed to study purely dissortative mixing between women and men in the population (i.e., only heterosexual HIV spread). This was a simplifying assumption of the model, since the RCT was targeted at women. While some of the HIV epidemic in the underlying population may be driven by male-to-male sex, there were no data on this available for this study population. In the model, we parameterized heterosexual mixing as an offset term for the `nodematch` network statistic for the attribute of sex. This involved fixing the coefficient at a value of negative infinity so that the term was not estimated in the model fitting procedures. Simulations from the model fit were consistent with this constraint.
3. **Degree distribution.** Without degree terms in the network model, the degree distribution (the number of persons with a momentary degree of 0, 1, 2, and so on) would follow a Poisson distribution with the rate parameter a function of the mean degree. In the Fem-PrEP data, there were 26% of women reporting secondary partners at baseline so we used this for the target statistic. This percentile did not vary by sex in our model, although men typically report much greater levels of concurrency than women in SSA [8]. This percentile was transformed into a number of concurrent nodes given the population sizes of men and women, and modeled using the `concurrent` statistic in ERGM.
4. **Sex-structured age homophily.** Typical with heterosexual mixing [9], our sub-Saharan African was modeled with a tendency to select sexual partners similar in age (age homophily). To model this, age was included as a continuous attribute with values following the empirical data. A new ERGM term was coded using the `ergm.userterms` software package to model the absolute difference in ages [10]. This network statistic includes a sum of the absolute difference that controls the variance of the distribution. Model diagnostics were checked to ensure that the mean and variance were preserved over changing age distributions that reflected demographic trends.

Dissolution model. In contrast to the formation model, the dissolution model includes only one network statistic, which is a function of the average duration of all partnerships. There have been several methods used to parameterize this statistic [11]. In any approach, the goal is to estimate

a coefficient for the dissolution model prior to estimating the formation model. It is therefore entered as a fixed offset term that specifies the average duration. This duration coefficient was used as the calibration parameter in our epidemic model. In order to fit the model results to HIV incidence observed in Fem-PrEP, the coefficient was adjusted until the baseline incidence rate in the age-eligible women was 5.0, similar to the baseline incidence of Fem-PrEP. That corresponded to a mean partnership duration of 380 days.

3 DEMOGRAPHY

In this model, there are three main demographic processes: entries, exits, and aging. Entries and exits are with respect to the sexually active population at which point persons become at risk of infection via heterosexual transmission, and so are modeled as starting at an age after birth and potentially ending at an age before death.

3.1 Entry/Birth

All persons enter the network at age 18, which was the lower age bound of the Fem-Prep study eligibility criteria [7]. The number of new entries at each time step is a proportional function of the current population size at that time step. This entry rate parameter was fixed so that there was a stable average growth rate in the population to maintain the stationary population distributions set at the outset of the simulation (25,000 persons). The number of entries into the network at each time step was simulated by drawing from a Poisson distribution with the rate parameter equal to the fixed entry rate times the current population size.

Incoming nodes were randomly assigned a sex and, for males, a circumcision status, among other relevant attributes. The probability of assignment as male sex was based on a draw from a binomial distribution with probability parameter set to maintain the sex ratio distribution at the outset of the simulation. This sex distribution was 50% male and 50% female. The specified level of circumcision in the baseline model was set to 40%, on the high end of estimates for the region from which the RCTs occurred [12]. The circumcision status was also randomly assigned at entry with by drawing from a binomial distribution with probability set to the specified level.

3.2 Exit/Mortality

All persons exit the network by age 55, either from death or cessation of sexual activity. The upper limit of age 55 was enforced deterministically for everyone in the network. The exits due to death were stochastic processes based on natural (non-HIV) and HIV-related causes before that age. Background mortality was modeled with age-specific and sex-specific mortality rates derived from the World Health Organization’s demographic life tables for Kenya, where one of the study sites was located [13].

Life table data provide the probability of death by age and sex for the population for the year 2010, closest to the date of the RCT. The mortality rates were then applied to active persons within the network through a stochastic process by drawing from a binomial series for each eligible person with a probability corresponding to that person's age and sex. Disease-related mortality was modeled based on clinical disease progression, and is described in Section 5 below.

3.3 Aging

The aging process in the population was linear in time for all active nodes. Nodes who exited the network from death were no longer active and their attributes were no longer updated. The unit of time step in these simulations was one week, and therefore, nodes were aged in weekly steps between the minimum and maximum ages allowed (18 and 55 years old).

4 INTRAHOST EPIDEMIOLOGY

Intrahost epidemiology includes those processes related to disease progression within infected persons. The two main components of progression within the HIV context are CD4 count and HIV viral load [14]. These clinical data come from large external cohort studies and clinical trials.

4.1 CD4 Progression

Persons have a CD4 count assigned at infection. Post-infection, that count declines naturally in the absence of anti-retroviral therapy (ART) and rises with successful ART. The underlying model for the decline follows Pantazis and colleagues, summarizing a meta-cohort of Africans [15]. This model has a non-linear decline in CD4 count based on age at infection and sex within Sub-Saharan Africa. Women have a higher baseline CD4 compared to men, and the rate of decline is faster when persons are infected at an older age. The decline is non-linear as the slopes are calculated on the scale of square roots of CD4 count.

Table S1 shows example values of the base CD4 count at infection and then the years it takes to reach certain threshold values of 350, 200, and 100 cells. As noted, women have a higher base value and progress to AIDS more slowly than men. Persons infected at younger ages will progress more slowly than those infected at older ages across sex.

Table S1. CD4 Clinical Model

	Base Value	Years to Threshold		
	Base CD4	*350*	*200*	*100*
Males				
Age Infected				
25	518.4	3.53	7.50	11.10
35	518.4	3.24	6.90	10.22
45	518.4	2.73	5.79	8.57
Females				
Age Infected				
25	570.3	4.50	8.47	12.07
35	570.3	4.13	7.79	11.10
45	570.3	3.47	6.54	9.31

Disease-induced mortality is a stochastic process when the CD4 crosses a late-stage AIDS threshold. Similar to Eaton and Hallett [16], we model this threshold level at a count of 50 CD4. Once persons hit that threshold, disease-related death occurs on average within one year. This is parameterized by a series of binomial draws with probability equal to an expected mortality in one year. Since this is a stochastic process of geometrically distributed waiting times, persons may randomly die before they reach a CD4 of 0, but that lower limited is enforced as a deterministic process.

4.2 HIV Viral Load

Following prior approaches [11], we parameterize changes in HIV viral load to account for the heighted viremia during acute-stage infection [17], viral set point during the long chronic stage of infection, and subsequent rise of viral load at clinical AIDS towards disease-related mortality. A starting viral load of 0 is assigned to all persons upon infection. From there, the natural viral load curve is fit with the following parameters.

Post-infection, it takes 14 days to reach peak viremia [18], at a level of 6.7 $\log_{10}$ [18, 19]. From peak viremia, it takes 107 days to reach viral set point [20], which is at a 4.5 $\log_{10}$ level [18, 19]. The total time of acute to chronic-stage infection will depend on CD4 progression using the model specified above. Rise of viral load begins during AIDS-stage infection, which occurs when the CD4 reaches a threshold of 200. This again follows the Pantazis disease progress model [15]. For men infected at age 45, that will occur 5.8 years after infection; for women infected at age 25, it will occur after 8.5 years.

During this late-stage infection period, viral load will rise from the set point of 4.5 $\log_{10}$ to a maximum of 7 $\log_{10}$ [14]. The time it takes to reach this maximum is also a function of the CD4 slope, and with the time it takes to transition from a CD4 of 200 to a CD4 of 50, at which point

late-stage AIDS with a high probability of mortality will occur [16]. These transitions are deterministic for all individuals. The VL trajectory is for ART-naive persons. The influence of ART on this trajectory is described below.

5 CLINICAL EPIDEMIOLOGY

Clinical epidemiological processes refer to disease diagnosis and treatment. In our model, diagnosis and treatment occur simultaneously based on the CD4 level of the respondent. This follows the empirical research of Collini et al [21] for one cohort in SSA that provides the mean and standard deviations of CD4 upon treatment initiation, which were 120 and 88 respectively.

5.1 Anti-Retroviral Therapy (ART) Initiation

Upon infection, each person is assigned a CD4 count at which they will initiate ART. This is drawn from a negative binomial distribution with mean and standard distribution parameters specified as stated above. Based on the statistical tests used, that study implied that the values were normally distributed. However, it was necessary to re-express this as distributed as a negative binomial because the parameter values (120 mean; 88 SD) result in 9% of the probability density below 0. We further scaled the standard deviation parameter for the negative binomial to 40 because the original value resulted in a tail where a considerable portion of the mass was above a threshold initiation value.

That threshold value for this study was based on an estimate for the time period in which the RCT occurred: persons may begin treatment at or below a CD4 of 350. Therefore, the draw for the individual-level count of when to initiate ART was capped at this 350 CD4 count value. Individuals could not initiate therapy above this count, and the average count at ART initialization followed the mean parameters.

There was an overall treatment coverage enforced among those reaching below their threshold values. Treatment coverage was expressed as the proportion of treatment-eligible persons (those with a CD4 under 350) who had ever initiated treatment. This was conservatively set at 25% in this population. Persons were allowed to initiate treatment, therefore, only if their current CD4 level had reach their targeted initiation value and less than 25% of eligible persons were currently “on ART,” even if their adherence to ART had lapsed.

5.2 ART Adherence

Adherence to ART was modeled to achieve an overall level of viral suppression among those who had ever initiated ART. Since ART adherence over the disease history is a complex dynamic process, with much uncertainty in how persons adhere to ART across disease stages, we used a calibration method in which we targeted that 80% of those who had initiated ART were virally

suppressed. This conforms to a recent UNAIDS analysis of viral suppression levels across Sub-Saharan Africa [22].

Modeling adherence in this way involved partitioning the on-treatment population into two groups: full adherers and partial adherers. Based on a meta-analysis of ART adherence that included many Sub-Saharan African countries, we set the full adherence proportion to 76% of the on-ART population [23]. Upon ART initiation, persons were assigned to the full versus partial adherence groups using a series of binomial draws with probability parameters set to this 76% value.

Full adherers remained on ART for the duration of their disease with no lapses. Sustained ART use conferred two clinical benefits: increased CD4 and decreased HIV viral load. Based on cohort data [21], the rebound in CD4 was modeled as a linear daily increase in CD4 level, with the slope conditional on sex of the person. CD4 count for males increased at an average of 9.75 per month; for females it was 11.6 CD4 cells per month. For full adherers, the CD4 increased back up to the level at initial infection, where it stayed for the duration of disease. This implied no disease-related mortality for those fully adherent on ART. ART was associated with a sustained decrease in HIV viral load within three months of initiation [24, 25], from the person's current level to viral suppression, which we specified at 1.5 $\log_{10}$, just below the limits of detection (1.65 $\log_{10}$) [11, 26].

Partial adherers cycled back and forth between ART use and non-use as a Markov process at each time step. The probability of cycling was calibrated to 50% at each time step, which resulted in the overall level of viral suppression of 80% among those who had initiated ART [22]. As a simplifying assumption and in the absence of robust data on rebound rates in western Africa, cycling off treatment resulted the reverse immunological and virological phenomena as cycling on treatment: CD4 declined at the negative rate it increased on ART and viral load increased at the positive rate it decreased on ART.

6 INTERHOST EPIDEMIOLOGY

Interhost epidemiological processes simulate the HIV-1 disease transmission in this model. Disease transmission occurs between sexual partners who are active on a given time step. The transmission rate formally provided below is expressed as a rate per partnership per unit time, which was one week in our model. This section will describe how the overall rate as a function of the intrahost epidemiological profile of each member of a partnership, and behavioral features within the dyad.

6.1 Disease-Discordant Dyads

At each time step in the simulation, a list of active dyads is selected from among the nodes in the network. This "edgelist" reflects the work of the network model simulation, wherein partnerships form on the basis of nodal attributes like age and sex. Dyads are active at a specific time step if terminus of that simulated edge is less than or equal to the current time step (right-censored). From this total edgelist, a disease-discordant subset is created by removing dyads in which both members are susceptible or both infected. This leaves those dyads that are serodiscordant on HIV status (one susceptible and one infected partner). This is the set of potential partnerships over which disease may be spread within that time step.

6.2 Per-Act Transmission Probability

Given an active serodiscordant dyad selected from an edgelist, next needed is a per-act transmission probability within that dyad. We use a probability model based on a recent study by Hughes and colleagues on transmissions within serodiscordant couples in chronic-stage infection [27]. Their statistical model relates the per-act transmission probability to the HIV viral load of the infecting partner, as well as the sex, age, and circumcision status of the susceptible partner. Also incorporated in their full model are co-infection with other sexually transmitted infections like Herpes Simplex Virus-2 (HSV-2). Because we did not have individual-level data on STIs, we used the population-level means assumed in the Hughes study.

The core component of the statistical model is the non-linear relationship between the HIV viral load of the infecting partner and the transmission probability given a sexual act. The association is linear in $\log_{10}$ VL, and therefore non-linear with respect to the outcome probability. The probability of transmission from an infected male partner with a $\log_{10}$ VL of 2, 3, 4, 5, 6, for example, to a susceptible female partner aged 35 would be: 0.0001, 0.0002, 0.0005, 0.0015, and 0.0044. The full model is parameterized in the *EpiModelHIV* function `hughes_tp` available on Github (LINK).

As noted, Hughes et al estimated transmission probabilities only for chronic and late-stage infection, and therefore their model does not account for the increased risk of transmission during acute-stage infection above and beyond that predicted by the peak viremia during that period [17]. Similar to Eaton and Hallet [16] and others modeling primary infection, we added a base multiplier of 5 onto the transmission probability if the infected partner was in the acute stage of infection. This brings the overall transmission probability during this stage in line with earlier predictions from Wawer [28] and Hollingsworth [29], which were better able to quantify the transmission probabilities during early infection.

6.3 Coital Frequency

Coital frequency is the number of acts of sexual intercourse per unit time. This behavioral element has been shown to vary widely by demographic and geographic groups [30]. This model uses coital rates from the baseline conditions in the Fem-PrEP model, which are outlined in Table 1 of the main paper. When translated into a weekly time step, this corresponds to 3.42 acts per week between active partners. Following data from Rakai, we assumed that this rate declined by half for partnerships in which the infected partner was in AIDS-stage disease [31]. With these rates, the total number of acts per partnership for that time step were then drawn from a Poisson distribution with the rate parameter set to that respective rate based on the disease stage of the infected partner.

6.4 Condom Use

Condom use is typically ignored or conflated into the overall act rate, but we separated it out based on the Fem-PrEP study baseline conditions that 51% of acts were protected by condoms. To generate a number of protected and unprotected acts at each time step for each serodiscordant dyad, we drew from a Poisson distribution for the total number of acts, and then from a binomial series for each act with a probability equal to the observed condom use. For the effect on transmission risk, we modeled condom use during sex associated with a 78% reduction in the risk of transmission per act, also derived from the Hughes per-act transmission probability model [27].

6.5 Pre-Exposure Prophylaxis

The allocation, uptake, and adherence to Pre-Exposure Prophylaxis (PrEP) are described in Section 8 below. Here we describe the impact of PrEP on the probability of HIV transmission given its use at any time step. Following estimates from a systematic review of PrEP trials, we implemented PrEP's effects on disease transmission by reducing the per-act transmission probability by a specified level estimated from per-protocol analyses from trial data [32]. We started with an assumed efficacy of 80% in baseline models, but varied this in counterfactual scenarios.

6.6 Transmission Rate

Given the elements above, for the persons not on PrEP the transmission rate per partnership per unit time is expressed as:

$$1 - [((1 - \tau(1 - \omega))^{\alpha_p}) \times ((1 - \tau)^{\alpha_u})]$$

where τ is the base transmission probability, ω is the efficacy of condoms, α_p is the number of acts protected by condoms, and α_u is the number of unprotected acts. Therefore, the overall

transmission rate is the 1 minus the product of probability of remaining susceptible due to protected acts and the probability of remaining susceptible due to unprotected acts.

For persons on PrEP, the transmission rate incorporates an additional element for the efficacy of PrEP:

$$1 - [((1 - \tau\beta(1 - \omega))^{\alpha_p}) \times ((1 - \tau\beta)^{\alpha_u})]$$

where β is the relative reduction in risk associated with PrEP use. The total transmission rate for the population was a weighted average of these two transmission rates.

7 RANDOMIZED CLINICAL TRIAL

The RCT within this model was simulated by recruiting women into the trial, randomly allocating them to a treatment or control arm, and following their time on trial until they reached the study end point, either through seroconversion or censoring. The specific components of that process are as follows.

First, we established the eligibility criteria for the study, which followed the model of Fem-PrEP, which enrolled women between the ages of 18 and 35. This was the only eligibility criterion in our model because the risk-related attributes were common to the entire underlying population. The simulated trial began at week three of the simulation. From the sampling frame of eligible women at each time step, women were randomly recruited from that list with no selection bias. The sampling frame could grow or shrink over time as women joined the trial and entered into the age-eligibility window.

The recruitment rate was set such that the recruitment was completed within two years of the initial recruitment date. The maximum recruitment time frame was longer than Fem-PrEP and was instead based on the VOICE study because Fem-PrEP was stopped early due to futility, and we sought to estimate effects unimpeded by a data safety monitoring board. Recruitment continued until the sample size reached a pre-defined coverage fraction. The coverage fraction represents the proportion of eligible persons on the sampling frame that will eventually be recruited into the RCT. The baseline coverage is 20%, but the counterfactual scenarios vary this from 10% to 50%.

After women are recruited into the trial, they are immediately randomized to the treatment or control arms in which they receive PrEP or a placebo, respectively. The randomization fraction is set to 50%, similar to Fem-PrEP. After randomization, women immediately initiate their treatment course. Women on the PrEP arm are randomly assigned an adherence type corresponding to full or partial adherence, similar to the ART adherence model outlined above. Full adherers stay on study drug for the entirety of their time on trial, while partial adherers cycle on and off dug as a Markov process at each time step. The levels of adherence were calibrated

such that the overall level of observable adherence stochastically varied around pre-defined levels of 25%, 50%, 75%, and 100%. The observable adherence was calculated similar to recent trials, as the proportion of study subjects with PrEP in their system at their study visit. The lowest level of adherence, 25% was similar to the observed levels of adherence in both Fem-PrEP and VOICE, and serves as a primary comparison in the results to validate the structure of the model.

Persons on trial were followed-up at monthly intervals for diagnostic HIV testing, following the intervals used in Fem-PrEP. The primary study endpoint was seroconversion documented at one of these monthly visits. Persons could also exit the trial through three modes of censoring: loss to follow-up, mortality, and maximum follow-up time. Persons were randomly lost to follow up at a rate matching the Fem-PrEP trial. Deaths rarely occurred due to the young age of the participants and their corresponding low mortality rate. Reaching the maximum follow-up time, corresponding to 36 months of observation, was the most common reason for censoring. This maximum follow-up time, similar to recruitment window, was based on the VOICE trial rather than the Fem-PrEP trial to ensure that the estimation of effects was not subject to early stoppage rules that occurred in Fem-PrEP.

The final simulated study data from each simulation contains the start and end times on trial, and the study outcome. This allowed for a standard survival analysis of data, which we conducted using both non-parametric Kaplan Meier methods and estimation of incidence rates under the standard Poisson distributional assumptions. For the data analysis, direct effects were estimated as the complement to the hazard ratio (1 – HR) comparing incidence in the treatment and placebo arms in the trial population, while the indirect effects was the similar expression of the comparison between incidence of the community control against the placebo arm in the trial population. The total effects of PrEP were therefore calculated as one minus the hazard ratio comparing community control against the treatment arm in the trial population.

8 INITIAL CONDITIONS AND MODEL CONFIGURATIONS

To initialize the epidemic simulations, it was necessary to set initial conditions of demography and epidemiology from which the natural infectious disease dynamics would flow. For our baseline models, we initialized the prevalence at 1% in order to generate an endemic HIV incidence rate matching the observed data. Similar to our prior models, a burn-in model was used to generate a population with stationary distributions of demographics, behavior, and disease incidence.

To start this burn-in simulation, we initialized epidemiological features of the population related to disease transmission. The time of infection was assigned by drawing from a geometric distribution with probability parameter equal to the average treatment-naive lifespan (roughly 12

years). Initial CD4 count and HIV viral load for those infected were based off of the clinical model described above, with the main parameters determining the disease stage linked to age at infection and sex. All infected persons were initialized as undiagnosed and untreated infection, but persons meeting their individualized CD4 threshold for ART were then immediately started upon the simulation.

The time unit for all simulations was one week. The choice of unit in microsimulation models like these is a tradeoff between computational efficiency (longer time units takes less computation given a fixed calendar time period for simulation) and competing risks. The latter is a concern for discrete-time dynamic models, as multiple processes must occur “at” a specific time. The discrete-time representation is an approximation to continuous time, in which individual events occur at distinct times [33]. In our models, a one-week interval for simulation provided the best balance between the bias of competing risks and computational efficiency. Smaller-scale diagnostic simulations at daily time step intervals (not shown) did not yield different substantive outcomes.

This burn-in model was simulated over 100 years of time, but reach a stable equilibrium point after approximately 50 years when the disease-related mortality associated with greater incidence had stabled and the age distribution was settled. As noted above, this burn-in period was used to calibrate the incidence at equilibrium to 5 per 100 person-years in the age-eligible subgroup of women who would be potential RCT study subjects. Matching to that level was based on a qualitative assessment of the stability in HIV incidence across the final 20 years of simulated time, where the deviation from observed incidence rates was less than 0.1% across that observation window.

At the conclusion of the burn-in period, which consisted of 100 simulations of the network, we selected that simulation which had the mean level of disease incidence across all simulations as a starting point for our intervention simulations. The first model that we conducted was the community control simulation. In this model, we simulated the RCT process as in all subsequent scenarios, but set the treatment efficacy of PrEP to 0. This allowed for a monitoring of the background incidence in the specific age-eligible group of women who would have participated in an effective trial. Similar to the RCT simulations that assumed an active agent, we conducted a total of 250 simulations across the same 36-month observation window. Incidence was calculated using the same survival methods.

CONCLUSION

This dissertation investigated how the unique configurations of sexual networks among heterosexuals in Sub-Saharan Africa (SSA) impacted two of the most effective biomedical tools for the prevention of HIV-1 infection. In this section, we highlight the major findings across three studies and suggest directions for future research.

Stochastic Network Modeling

The power of mathematical modeling for epidemics in general is its implementation science and population health framework, which can help translate findings from clinical trials and observational studies to contribute to public health policy.

The model work in this dissertation is based on a relatively new approach to the mathematical modeling for infectious disease. Our network models are grounded in the statistical framework of temporal exponential-family random graph models (ERGMs), with which one may simulate a complete network from egocentrically collected data from a population-based survey as we described in Chapter 1. These models allow for the representation of sexual partnership dynamics in a way that the prevailing modeling methods (deterministic compartmental models) do not. Accordingly, they require fewer assumptions about extreme high-risk behaviors to generate epidemics consistent with those observed in SSA. Other benefits of this approach are similar to microsimulation models in general: the representation of continuous attributes like HIV viral load and age; and the assessment of the stochastic variability in outcomes given a particular set of parameters.

The research applications in Chapters 2 and 3 of this dissertation have benefitted from and coincided with the implementation of these methods into a free, open-source software tool, *EpiModel*, developed and maintained by our research team. *EpiModel* (www.epimodel.org) provides a generalized framework for the mathematical modeling of infectious disease, with a programming interface that facilitates model extensions for novel research applications.

Our software is under continual and extensive development. Future development work will include the release of an extension software package, *EpiModelHIV*, featuring many of the methods presented here. This extension package will start with heterosexual transmission in the SSA context, but will grow to support the network modeling of HIV in different populations, including men who have sex with men and injection drug users. This software development effort will provide the foundation for other researchers to use these powerful new methods to answer key questions related to the population-level impact of new HIV interventions.

Rapid Primary Data Collection to Support Modeling

Often mathematical modeling is done in a vacuum, with no connections to the modeled populations, and parameterized through a collection of secondary data of questionable representativeness. The mathematical models in this project were an extension of our initial empirical research, the Migration & Health in Ghana study. Our research group developed the MHG survey tool with the goal of dynamic modeling in mind, and thus we included questions specifically for parameterizing our models. Modeling should be supported by targeted primary data collection when necessary to maintain the link to the target population. The initial conception of the models on circumcision and concurrency in Chapter 2 was to incorporate coital dilution into the models. However, our analysis of empirical data in Chapter 1, which found essentially no dilution, motivating our decision to not include it in those models.

In the long term, if the current research funding climate continues to limit expensive longitudinal studies, integrated brief data collection with mathematical modeling provides an alternative approach to gain scientific insight into "big" questions but in a cost-effective manner. Integrating modeling into traditional epidemiological research has been a growing priority: the MP3 funding mechanism at NIH, which funds studies to test the feasibility of novel combination prevention packages for HIV, requires a mathematical modeling component to complement empirical research methods. Our PreP model in Chapter 3 suggests an even more integrated approach at an earlier stage: that study provides a framework for making decisions about the study design and setting. Our future research in this area will consider how modeling can specifically inform clinical trials at all stages: design, interim analysis, and interpretation of study findings in the trial itself, along with more common forecasting at the population level.

Steep Thresholds for Network Connectivity

A main finding from our circumcision model in Chapter 2 was that changing concurrency levels by 10% in absolute terms conferred greater reductions in HIV incidence than changing circumcision by 80%. This 8-fold relative difference in effect magnitude stems from the strong non-linear relationships between network connectivity and disease transmission. Individual biological changes like circumcision also yield non-linearity (i.e., greater than 1 infection is averted per circumcision performed) but the relationship is more closely linear than for network phenomena. In Chapter 3, we found that reducing the network connectivity by half resulted in a reduction in baseline incidence by more than half, and this population scenario presented a better environment for the estimation of PreP efficacy.

Although no RCTs have ever found that behavior change alone has resulted in decreased HIV incidence, all proven biomedical tools have a behavioral component (e.g., drug adherence), and trials of novel prevention agents always provide a behavioral risk reduction element (e.g., ongoing counseling at study visits). The widespread variability in sexual risk behaviors across SSA suggests that they are not biological determined attributes, but modifiable traits. Our research suggests that network-related

behaviors like concurrency and sex-structured age homophily should also be a focus of behavior change. Network concepts could be integrated into combination prevention packages that implement circumcision, PreP, or ART with modest behavioral change to yield significant reductions in disease incidence. Our Chapter 1 study suggests that modest levels of risk reduction related to condom use is being practiced with those with a high level of concurrency (3+ partners at any time); while total per-partnership acts did not decline during these high-risk times, the unprotected acts did modestly. This evidence for selective condom use could be the basis of an effective intervention if paired with biomedical tools.

Behavioral Targeting of Interventions and Research

One theme common to this dissertation was the network context matters. It matters when forecasting the effectiveness of a prevention tool at the population level, and when estimating the efficacy of a tool with a study setting. Increases in circumcision had a greater relative effect on HIV incidence in networks with low connectivity. More highly connected networks led to a larger bias to the estimation of PreP efficacy, compared to less connected networks where the baseline disease incidence was lower. Even in the absence of behavioral change, this dissertation provides evidence that behavioral targeting of HIV prevention activities and research should start with the network context. This targeting can provide a cost-effective approach to investigating and implementing effective HIV prevention interventions.

REFERENCES

Chapter 1: Minimal Coital Dilution in Accra, Ghana

1. 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.
2. Morris M, Kretzschmar M. Concurrent partnerships and the spread of HIV. Aids. 1997;11(5):641-8.
3. Goodreau SM, Cassels S, Kasprzyk D, Montano DE, Greek A, Morris M. Concurrent partnerships, acute infection and HIV epidemic dynamics among young adults in Zimbabwe. AIDS and behavior. 2012;16(2):312-22.
4. Reniers G, Tfaily R. Polygyny, partnership concurrency, and HIV transmission in Sub-Saharan Africa. Demography. 2012;49(3):1075-101.
5. Gaydosh L, Reniers G, Helleringer S. Partnership Concurrency and Coital Frequency. AIDS and behavior. 2013;17(7):2376-86.
6. 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.
7. Weller S, Davis K. Condom effectiveness in reducing heterosexual HIV transmission. The Cochrane database of systematic reviews. 2002;(1):CD003255.
8. Delva W, Meng F, Beauclair R, et al. Coital frequency and condom use in monogamous and concurrent sexual relationships in Cape Town, South Africa. Journal of the International AIDS Society. 2013;16:18034.
9. Johnson AM, Mercer CH, Erens B, et al. Sexual behaviour in Britain: partnerships, practices, and HIV risk behaviours. Lancet. 2001;358(9296):1835-42.
10. Oberhauser AM, Yeboah MA. Heavy burdens: Gendered livelihood strategies of porters in Accra, Ghana. Singapore Journal of Tropical Geography. 2011;32(1):22-37.
11. HIV/AIDS JUNPo. Global report: UNAIDS report on the global AIDS epidemic: 2012: UNAIDS; 2012.
12. Cassels S, Jenness SM, Biney A, Ampofo W, Dodoo F. Migration, Sexual Networks, and HIV in Ghana: Study Design and Preliminary Results. Under review.
13. Luke N, Clark S, Zulu EM. The relationship history calendar: improving the scope and quality of data on youth sexual behavior. Demography. 2011;48(3):1151-76.
14. Lumley T. Complex surveys: A guide to analysis using R: John Wiley & Sons; 2011.
15. Imai K, King G, Lau O. Zelig: Everyone's statistical software. R package version. 2009;3(5).
16. Rothman KJ, Greenland S, Lash TL. Modern epidemiology: Wolters Kluwer Health; 2008.
17. Bailey RC, Moses S, Parker CB, et al. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. Lancet. 2007;369(9562):643-56.
18. Gray RH, Kigozi G, Serwadda D, et al. Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. Lancet. 2007;369(9562):657-66.
19. 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.

Chapter 2: Effectiveness of Male Circumcision for HIV-1 Prevention Depends on Contact Network Structure

Main Study

1. Kurth AE, Celum C, Baeten JM, Vermund SH, Wasserheit JN. Combination HIV prevention: significance, challenges, and opportunities. Current HIV/AIDS reports. 2011;8(1):62-72.
2. Padian NS, Buve A, Balkus J, Serwadda D, Cates W, Jr. Biomedical interventions to prevent HIV infection: evidence, challenges, and way forward. Lancet. 2008;372(9638):585-99.

3. 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):e84511.
4. Dickson KE, Tran NT, Samuelson JL, et al. Voluntary medical male circumcision: a framework analysis of policy and program implementation in eastern and southern Africa. PLoS medicine. 2011;8(11):e1001133.
5. Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, Puren A. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: the ANRS 1265 Trial. PLoS medicine. 2005;2(11):e298.
6. Bailey RC, Moses S, Parker CB, et al. Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. Lancet. 2007;369(9562):643-56.
7. Gray RH, Kigozi G, Serwadda D, et al. Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. Lancet. 2007;369(9562):657-66.
8. Wawer MJ, Makumbi F, Kigozi G, et al. Circumcision in HIV-infected men and its effect on HIV transmission to female partners in Rakai, Uganda: a randomised controlled trial. Lancet. 2009;374(9685):229-37.
9. Hallett TB, Alsallaq RA, Baeten JM, et al. Will circumcision provide even more protection from HIV to women and men? New estimates of the population impact of circumcision interventions. Sexually transmitted infections. 2011;87(2):88-93.
10. Alsallaq RA, Cash B, Weiss HA, et al. Quantitative assessment of the role of male circumcision in HIV epidemiology at the population level. Epidemics. 2009;1(3):139-52.
11. Halperin DT, Epstein H. Why is HIV prevalence so severe in southern Africa?: the role of multiple concurrent partnerships and lack of male circumcision-implications for HIV prevention: opinion. Southern African Journal of HIV Medicine. 2007;(26):19-23, 5.
12. Smith RJ, Li J, Gordon R, Heffernan JM. Can we spend our way out of the AIDS epidemic? A world halting AIDS model. BMC public health. 2009;9 Suppl 1:S15.
13. Hallett TB, Singh K, Smith JA, White RG, Abu-Raddad LJ, Garnett GP. Understanding the impact of male circumcision interventions on the spread of HIV in southern Africa. PloS one. 2008;3(5):e2212.
14. Alsallaq RA, Baeten JM, Celum CL, et al. Understanding the potential impact of a combination HIV prevention intervention in a hyper-endemic community. PloS one. 2013;8(1):e54575.
15. Morris M, Kretzschmar M. Concurrent partnerships and the spread of HIV. Aids. 1997;11(5):641-8.
16. Cassels S, Jenness SM, Biney AA, Ampofo WK, Dodoo FN. Migration, sexual networks, and HIV in Agbogbloshie, Ghana. Demographic research. 2014;31:861-88.
17. 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.
18. 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.
19. 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.
20. Pantazis N, Morrison C, Amornkul PN, et al. Differences in HIV natural history among African and non-African seroconverters in Europe and seroconverters in sub-Saharan Africa. PloS one. 2012;7(3):e32369.
21. Little SJ, McLean AR, Spina CA, Richman DD, Havlir DV. Viral dynamics of acute HIV-1 infection. The Journal of experimental medicine. 1999;190(6):841-50.
22. Mellors JW, Munoz A, Giorgi JV, et al. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. Annals of internal medicine. 1997;126(12):946-54.

23. Morrison CS, Chen PL, Nankya I, et al. Hormonal contraceptive use and HIV disease progression among women in Uganda and Zimbabwe. Journal of acquired immune deficiency syndromes. 2011;57(2):157-64.
24. Pilcher CD, Price MA, Hoffman IF, et al. Frequent detection of acute primary HIV infection in men in Malawi. Aids. 2004;18(3):517-24.
25. Collini P, Schwab U, Sarfo S, et al. Sustained immunological responses to highly active antiretroviral therapy at 36 months in a Ghanaian HIV cohort. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2009;48(7):988-91.
26. Ampofo WK. Current Status of HIV/AIDS Treatment, Care and Support Services in Ghana. Ghana medical journal. 2009;43(4):142-3.
27. HIV/AIDS JUNPo. The Gap Report. Geneva: UNAIDS. 2014.
28. Chu H, Gange SJ, Li X, et al. The effect of HAART on HIV RNA trajectory among treatment-naive men and women: a segmental Bernoulli/lognormal random effects model with left censoring. Epidemiology. 2010;21 Suppl 4:S25-34.
29. Goodreau SM, Carnegie NB, Vittinghoff E, et al. What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? PloS one. 2012;7(11):e50522.
30. Hughes JP, Baeten JM, Lingappa JR, et al. Determinants of per-coital-act HIV-1 infectivity among African HIV-1-serodiscordant couples. The Journal of infectious diseases. 2012;205(3):358-65.
31. 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.
32. Eaton JW, Hallett TB, Garnett GP. Concurrent sexual partnerships and primary HIV infection: a critical interaction. AIDS and behavior. 2011;15(4):687-92.
33. 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 and behavior. 2013.
34. Goodreau SM, Cassels S, Kasprzyk D, Montaño DE, Greek A, Morris M. Concurrent Partnerships, Acute Infection and HIV Epidemic Dynamics Among Young Adults in Zimbabwe. AIDS and behavior. 2010;16(2):312-22.
35. Mills EJ, Nachega JB, Buchan I, et al. Adherence to antiretroviral therapy in sub-Saharan Africa and North America: a meta-analysis. Jama. 2006;296(6):679-90.
36. Cassels S, Manhart L, Jenness SM, Morris M. Short-term Mobility and Increased Partnership Concurrency among Men in Zimbabwe. PloS one. 2013;8(6):e66342.
37. 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. Sexually transmitted diseases. 2012;39(6):407-13.
38. Padian NS, McCoy SI, Manian S, Wilson D, Schwartlander B, Bertozzi SM. Evaluation of large-scale combination HIV prevention programs: essential issues. Journal of acquired immune deficiency syndromes. 2011;58(2):e23-8.
39. Carnegie NB, Morris M. Size matters: concurrency and the epidemic potential of HIV in small networks. PloS one. 2012;7(8):e43048.
40. Helleringer S, Kohler HP, Chimbiri A, Chatonda P, Mkandawire J. The Likoma Network Study: Context, data collection, and initial results. Demographic research. 2009;21:427-68.
41. Cassels S, Jenness SM, Khanna AS. Conceptual framework and research methods for migration and HIV transmission dynamics. AIDS and behavior. 2014;18(12):2302-13.

Supplemental Appendix

1. Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M. statnet: Software Tools for the Representation, Visualization, Analysis and Simulation of Network Data. Journal of statistical software. 2008;24(1):1548-7660.
2. Lessler J, Edmunds WJ, Halloran M, Hollingsworth TD, Lloyd AL. Seven challenges for model-driven data collection in experimental and observational studies. Epidemics. 2014.
3. Cassels S, Jenness SM, Biney AA, Ampofo WK, Dodoo FN. Migration, sexual networks, and HIV in Agbogbloshie, Ghana. Demographic research. 2014;31:861-88.
4. Luke N, Clark S, Zulu EM. The relationship history calendar: improving the scope and quality of data on youth sexual behavior. Demography. 2011;48(3):1151-76.
5. 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.
6. van Duijn MA, Gile KJ, Handcock MS. A Framework for the Comparison of Maximum Pseudo Likelihood and Maximum Likelihood Estimation of Exponential Family Random Graph Models. Social networks. 2009;31(1):52-62.
7. 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.
8. Carnegie NB, Krivitsky PN, Hunter DR, Goodreau SM. An approximation method for improving dynamic network model fitting. Journal of Computational and Graphical Statistics. 2014;(just-accepted):00-.
9. Krivitsky PN, Handcock MS, Morris M. Adjusting for Network Size and Composition Effects in Exponential-Family Random Graph Models. Statistical methodology. 2011;8(4):319-39.
10. 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.
11. Hunter DR, Goodreau SM, Handcock MS. ergm.userterms: A Template Package for Extending statnet. Journal of statistical software. 2013;52(2):i02.
12. Goodreau SM, Carnegie NB, Vittinghoff E, et al. What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? PloS one. 2012;7(11):e50522.
13. Burington B, Hughes JP, Whittington WLH, et al. Estimating duration in partnership studies: issues, methods and examples. Sexually transmitted infections. 2010;86(2):84-9.
14. Cassels S, Jenness SM, Khanna AS. Conceptual framework and research methods for migration and HIV transmission dynamics. AIDS and behavior. 2014;18(12):2302-13.
15. Halperin DT, Epstein H. Why is HIV prevalence so severe in southern Africa?: the role of multiple concurrent partnerships and lack of male circumcision-implications for HIV prevention: opinion. Southern African Journal of HIV Medicine. 2007;(26):19-23, 5.
16. Organization WH. World health statistics 2010: World Health Organization; 2010.
17. Mellors JW, Munoz A, Giorgi JV, et al. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. Annals of internal medicine. 1997;126(12):946-54.
18. Pantazis N, Morrison C, Amornkul PN, et al. Differences in HIV natural history among African and non-African seroconverters in Europe and seroconverters in sub-Saharan Africa. PloS one. 2012;7(3):e32369.
19. 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.

20. Fiebig EW, Wright DJ, Rawal BD, et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. Aids. 2003;17(13):1871-9.
21. Little SJ, McLean AR, Spina CA, Richman DD, Havlir DV. Viral dynamics of acute HIV-1 infection. The Journal of experimental medicine. 1999;190(6):841-50.
22. Pilcher CD, Price MA, Hoffman IF, et al. Frequent detection of acute primary HIV infection in men in Malawi. Aids. 2004;18(3):517-24.
23. Morrison CS, Chen PL, Nankya I, et al. Hormonal contraceptive use and HIV disease progression among women in Uganda and Zimbabwe. Journal of acquired immune deficiency syndromes. 2011;57(2):157-64.
24. Collini P, Schwab U, Sarfo S, et al. Sustained immunological responses to highly active antiretroviral therapy at 36 months in a Ghanaian HIV cohort. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2009;48(7):988-91.
25. Ampofo WK. Current Status of HIV/AIDS Treatment, Care and Support Services in Ghana. Ghana medical journal. 2009;43(4):142-3.
26. HIV/AIDS JUNPo. The Gap Report. Geneva: UNAIDS. 2014.
27. Mills EJ, Nachega JB, Buchan I, et al. Adherence to antiretroviral therapy in sub-Saharan Africa and North America: a meta-analysis. Jama. 2006;296(6):679-90.
28. Chu H, Gange SJ, Li X, et al. The effect of HAART on HIV RNA trajectory among treatment-naive men and women: a segmental Bernoulli/lognormal random effects model with left censoring. Epidemiology. 2010;21 Suppl 4:S25-34.
29. Hammer SM, Eron JJ, Jr., Reiss P, et al. Antiretroviral treatment of adult HIV infection: 2008 recommendations of the International AIDS Society-USA panel. Jama. 2008;300(5):555-70.
30. Rizzardi GP, De Boer RJ, Hoover S, et al. Predicting the duration of antiviral treatment needed to suppress plasma HIV-1 RNA. The Journal of clinical investigation. 2000;105(6):777-82.
31. Hughes JP, Baeten JM, Lingappa JR, et al. Determinants of per-coital-act HIV-1 infectivity among African HIV-1-serodiscordant couples. The Journal of infectious diseases. 2012;205(3):358-65.
32. Wawer MJ, Gray RH, Sewankambo NK, et al. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. The Journal of infectious diseases. 2005;191(9):1403-9.
33. Hollingsworth TD, Anderson RM, Fraser C. HIV-1 transmission, by stage of infection. The Journal of infectious diseases. 2008;198(5):687-93.
34. 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.
35. 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.
36. Delva W, Meng F, Beauclair R, et al. Coital frequency and condom use in monogamous and concurrent sexual relationships in Cape Town, South Africa. Journal of the International AIDS Society. 2013;16:18034.
37. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals: Princeton University Press; 2008.
38. Abu-Raddad LJ, Longini IM, Jr. No HIV stage is dominant in driving the HIV epidemic in sub-Saharan Africa. Aids. 2008;22(9):1055-61.
39. Goodreau SM. Is 2 a "high number of partners"? Modeling, data, and the power of concurrency. Sexually transmitted diseases. 2013;40(1):61.
40. Alkema L, Raftery AE, Brown T. Bayesian melding for estimating uncertainty in national HIV prevalence estimates. Sexually transmitted infections. 2008;84 Suppl 1:i11-i6.
41. 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.

Chapter 3: Interference for HIV-1 Pre-Exposure Prophylaxis Efficacy in Dynamic Contact Networks

Main Study

1. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals: Princeton University Press; 2008.
2. Halloran ME, Struchiner CJ. Designs for Dependent Happenings. Epidemiology. 1991;2(5):331-8.
3. Anderson RM, May RM. Immunisation and herd immunity. Lancet. 1990;335(8690):641-5.
4. Halloran ME, Struchiner CJ. Causal inference in infectious diseases. Epidemiology. 1995;6(2):142-51.
5. Pearl J. Causality: models, reasoning and inference: Cambridge Univ Press; 2000.
6. Rubin DB. Causal inference using potential outcomes. J Am Stat Assoc. 2005;100(469):322-31.
7. 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):e84511.
8. Tchetgen EJT, VanderWeele TJ. On causal inference in the presence of interference. Statistical Methods in Medical Research. 2010;21(1):55-75.
9. Vanderweele TJ, Tchetgen Tchetgen EJ, Halloran ME. Components of the indirect effect in vaccine trials: identification of contagion and infectiousness effects. Epidemiology. 2012;23(5):751-61.
10. Halloran ME, Haber M, Longini IM, Jr., Struchiner CJ. Direct and indirect effects in vaccine efficacy and effectiveness. American journal of epidemiology. 1991;133(4):323-31.
11. Ogburn E, Vanderweele TJ. Causal Diagrams for Interference.
12. Baeten JM, Donnell D, Ndase P, et al. Antiretroviral Prophylaxis for HIV Prevention in Heterosexual Men and Women. N Engl J Med. 2012;367(5):399-410.
13. Grant RM, Lama JR, Anderson PL, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med. 2010;363(27):2587-99.
14. Choopanya K, Martin M, Suntharasamai P, et al. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet. 2013;381(9883):2083-90.
15. Van Damme L, Corneli A, Ahmed K, et al. Preexposure Prophylaxis for HIV Infection among African Women. N Engl J Med. 2012;367(5):411-22.
16. Marrazzo JM, Ramjee G, Richardson BA, et al. Tenofovir-based preexposure prophylaxis for HIV infection among African women. N Engl J Med. 2015;372(6):509-18.
17. Morris M, Kretzschmar M. Concurrent partnerships and the spread of HIV. Aids. 1997;11(5):641-8.
18. 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.
19. 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.
20. 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.
21. Burington B, Hughes JP, Whittington WLH, et al. Estimating duration in partnership studies: issues, methods and examples. Sex Transm Infect. 2010;86(2):84-9.
22. 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.
23. Goodreau SM, Carnegie NB, Vittinghoff E, et al. What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? PloS one. 2012;7(11):e50522.

24. Collini P, Schwab U, Sarfo S, et al. Sustained immunological responses to highly active antiretroviral therapy at 36 months in a Ghanaian HIV cohort. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2009;48(7):988-91.
25. HIV/AIDS JUNPo. The Gap Report. Geneva: UNAIDS. 2014.
26. Chu H, Gange SJ, Li X, et al. The effect of HAART on HIV RNA trajectory among treatment-naive men and women: a segmental Bernoulli/lognormal random effects model with left censoring. Epidemiology. 2010;21 Suppl 4:S25-34.
27. Hughes JP, Baeten JM, Lingappa JR, et al. Determinants of per-coital-act HIV-1 infectivity among African HIV-1-serodiscordant couples. The Journal of infectious diseases. 2012;205(3):358-65.
28. Baeten J, Celum C. Oral antiretroviral chemoprophylaxis. Curr Opin HIV AIDS. 2012;7(6):514-9.
29. Dai JY, Gilbert PB, Hughes JP, Brown ER. Estimating the efficacy of preexposure prophylaxis for HIV prevention among participants with a threshold level of drug concentration. American journal of epidemiology. 2013;177(3):256-63.
30. Hudgens MG, Halloran ME. Toward Causal Inference With Interference. J Am Stat Assoc. 2008;103(482):832-42.
31. Miller Neilan RL, Schaefer E, Gaff H, Fister KR, Lenhart S. Modeling optimal intervention strategies for cholera. Bulletin of mathematical biology. 2010;72(8):2004-18.
32. Alsallaq RA, Baeten JM, Celum CL, et al. Understanding the potential impact of a combination HIV prevention intervention in a hyper-endemic community. PloS one. 2013;8(1):e54575.
33. Boily MC, Lowndes CM, Vickerman P, et al. Evaluating large-scale HIV prevention interventions: study design for an integrated mathematical modelling approach. Sex Transm Infect. 2007;83(7):582-9.
34. Baeten JM, Heffron R, Kidoguchi L, et al. Near Elimination of HIV Transmission in a Demonstration Project of PreP and ART. Conference on Retroviruses and Opportunistic Infections; Seattle, WA2015.
35. Goodreau SM, Cassels S, Kasprzyk D, Montaño DE, Greek A, Morris M. Concurrent Partnerships, Acute Infection and HIV Epidemic Dynamics Among Young Adults in Zimbabwe. AIDS Behav. 2010;16(2):312-22.

Supplemental Appendix

1. Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M. statnet: Software Tools for the Representation, Visualization, Analysis and Simulation of Network Data. Journal of statistical software. 2008;24(1):1548-7660.
2. 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.
3. van Duijn MA, Gile KJ, Handcock MS. A Framework for the Comparison of Maximum Pseudo Likelihood and Maximum Likelihood Estimation of Exponential Family Random Graph Models. Social networks. 2009;31(1):52-62.
4. 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.
5. Carnegie NB, Krivitsky PN, Hunter DR, Goodreau SM. An approximation method for improving dynamic network model fitting. Journal of Computational and Graphical Statistics. 2014;(just-accepted):00-.
6. Krivitsky PN, Handcock MS, Morris M. Adjusting for Network Size and Composition Effects in Exponential-Family Random Graph Models. Statistical methodology. 2011;8(4):319-39.
7. Van Damme L, Corneli A, Ahmed K, et al. Preexposure Prophylaxis for HIV Infection among African Women. N Engl J Med. 2012;367(5):411-22.

8. 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. 2013.
9. 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.
10. Hunter DR, Goodreau SM, Handcock MS. ergm.userterms: A Template Package for Extending statnet. Journal of statistical software. 2013;52(2):i02.
11. Goodreau SM, Carnegie NB, Vittinghoff E, et al. What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)? PloS one. 2012;7(11):e50522.
12. Hallett TB, Alsallaq RA, Baeten JM, et al. Will circumcision provide even more protection from HIV to women and men? New estimates of the population impact of circumcision interventions. Sex Transm Infect. 2011;87(2):88-93.
13. Organization WH. World health statistics 2010: World Health Organization; 2010.
14. Mellors JW, Munoz A, Giorgi JV, et al. Plasma viral load and CD4+ lymphocytes as prognostic markers of HIV-1 infection. Annals of internal medicine. 1997;126(12):946-54.
15. Pantazis N, Morrison C, Amornkul PN, et al. Differences in HIV natural history among African and non-African seroconverters in Europe and seroconverters in sub-Saharan Africa. PloS one. 2012;7(3):e32369.
16. 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.
17. Fiebig EW, Wright DJ, Rawal BD, et al. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. Aids. 2003;17(13):1871-9.
18. Little SJ, McLean AR, Spina CA, Richman DD, Havlir DV. Viral dynamics of acute HIV-1 infection. The Journal of experimental medicine. 1999;190(6):841-50.
19. Pilcher CD, Price MA, Hoffman IF, et al. Frequent detection of acute primary HIV infection in men in Malawi. Aids. 2004;18(3):517-24.
20. Morrison CS, Chen PL, Nankya I, et al. Hormonal contraceptive use and HIV disease progression among women in Uganda and Zimbabwe. Journal of acquired immune deficiency syndromes. 2011;57(2):157-64.
21. Collini P, Schwab U, Sarfo S, et al. Sustained immunological responses to highly active antiretroviral therapy at 36 months in a Ghanaian HIV cohort. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2009;48(7):988-91.
22. HIV/AIDS JUNPo. The Gap Report. Geneva: UNAIDS. 2014.
23. Mills EJ, Nachega JB, Buchan I, et al. Adherence to antiretroviral therapy in sub-Saharan Africa and North America: a meta-analysis. Jama. 2006;296(6):679-90.
24. Chu H, Gange SJ, Li X, et al. The effect of HAART on HIV RNA trajectory among treatment-naive men and women: a segmental Bernoulli/lognormal random effects model with left censoring. Epidemiology. 2010;21 Suppl 4:S25-34.
25. Hammer SM, Eron JJ, Jr., Reiss P, et al. Antiretroviral treatment of adult HIV infection: 2008 recommendations of the International AIDS Society-USA panel. Jama. 2008;300(5):555-70.
26. Rizzardi GP, De Boer RJ, Hoover S, et al. Predicting the duration of antiviral treatment needed to suppress plasma HIV-1 RNA. The Journal of clinical investigation. 2000;105(6):777-82.
27. Hughes JP, Baeten JM, Lingappa JR, et al. Determinants of per-coital-act HIV-1 infectivity among African HIV-1-serodiscordant couples. The Journal of infectious diseases. 2012;205(3):358-65.

28. Wawer MJ, Gray RH, Sewankambo NK, et al. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. The Journal of infectious diseases. 2005;191(9):1403-9.
29. Hollingsworth TD, Anderson RM, Fraser C. HIV-1 transmission, by stage of infection. The Journal of infectious diseases. 2008;198(5):687-93.
30. 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.
31. Goodreau SM, Cassels S, Kasprzyk D, Montaño DE, Greek A, Morris M. Concurrent Partnerships, Acute Infection and HIV Epidemic Dynamics Among Young Adults in Zimbabwe. AIDS Behav. 2010;16(2):312-22.
32. Baeten J, Celum C. Oral antiretroviral chemoprophylaxis. Curr Opin HIV AIDS. 2012;7(6):514-9.
33. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals: Princeton University Press; 2008.