Title: Concurrent partnerships and the proportion of transmissions by stage of infection for HIV in Zimbabwe

Short title: Stage of infection for HIV in Zimbabwe

Authors: Steven M. GOODREAU, Susan CASSELS, Danuta KASPRZYK, Daniel E. MONTAÑO, April GREEK, Martina MORRIS

Corresponding author (incl. reprint requests):

Steven M. Goodreau

Department of Anthropology

Box 353100

University of Washington

Seattle, WA 98195

Word count*:* 3,500

Sources of funding: SMG and SC were supported in part by the Puget Sound Partners for Global Health (Research and Technology Project Award 26145). Data collection was supported by NIH (U10MH061544 and RZIAA014802). The models and software used in this study were developed with support from NIH (R01 HD041877 and R01 DA12831). SC was partially supported by NIH (K99 HD057553 and P30 AI27757).

ABSTRACT

Objective

To estimate the proportion of HIV infections stemming from acute stage cases using a model that incorporates information on partnership duration and overlap (concurrency)

Methods

We use a dynamic exponential random graph model (ERGM) to estimate partnership network parameters from data and drive a stochastic, dynamic simulation of HIV transmission through this network. Data come from a study of sexual behavior using egocentrically-sampled networks in thirty Zimbabwean communities. The model incorporates relational types (marital/live-in, steady, casual) by sex, concurrency by relational type, and relational duration by relational type. We conduct sensitivity analyses for the impact of concurrency and with four published estimates of transmission probabilities by time since infection based on data from Wawer et al. [1]

Results

Our models generally predict that around 20-25% of transmissions stem from acute-stage infections, 30-50% from chronic-stage, and 30-45% from AIDS-stage. The role of acute infections was more strongly affected by concurrency than by the range of published stage-specific transmission probabilities. If we constrain the model to allow only serial monogamy, as opposed to reported levels of concurrency, the simulated epidemic is unsustainable. The same is true if the acute transmission peak is eliminated.

Conclusions

Although acute transmission appears to account for a minority of new infections, reducing its impact -- by reducing either acute viral load or the probability someone has additional partners during early infection – could bring the epidemic below the reproductive threshold in populations marked by low rates of partner change but relatively high rates of long-term partner concurrency.

Key words: HIV/AIDS, mathematical modeling, concurrency, acute infection, sub-Saharan Africa, Zimbabwe

INTRODUCTION

HIV infection is defined by three general stages of disease – acute infection, chronic infection, and AIDS – coinciding with variation in viral load, CD4 cell count, and clinical manifestations [2-5]. Transmission probability also peaks during acute infection, decreases during chronic infection, and increases again during AIDS [1]. However, the population-level impact of this variation is not straightforward. The fraction of new infections stemming from each stage depends on multiple biological and behavioral processes, including relative transmission probabilities, stage duration, distribution of individuals by stage, and partnership dynamics.

For over a decade researchers have considered the proportion of transmission stemming from people within each stage [6-9]. However, solid empirical estimates of stage-specific per-coital act transmission probabilities were lacking until a recent landmark study [1]. Other investigators have since reevaluated the study data to develop new stage-specific per-act estimates [10-12]. Using these calculations, researchers have estimated the proportion of new HIV infections by stage [10-18]. Methods and results have varied widely, concluding that acute infections generate anywhere from 1-41% of infections in different populations (Table 1). Additional work, largely among homosexually-active men, has estimated the proportion using phylogenetics, concluding that nearly half of transmissions stems from acute-phase cases [19-21].

None of these studies explores a fully dynamic model incorporating behavioral data on relational overlap, although some have discussed the importance of doing so [12]. In order for someone to transmit while acutely infected, they must acquire the infection from one partner and transmit to

another during a brief window; i.e., they must have two partnerships close or overlapping in time. Previous modeling work has shown that partnership timing can have a major impact on the size and rapidity of HIV spread [22, 23], although this work did not incorporate stage-varying transmission. This paper integrates recent transmission estimates with data on partnership formation, dissolution and overlap from Zimbabwe using new methods that provide a unified framework for statistical estimation and epidemic simulation.

METHODS

Data derive from the pre-intervention baseline survey of sexually active adults in ZiCHIRe, the Zimbabwe arm of a five-nation popular opinion leader behavioral intervention study. The overall study and Zimbabwe-specific additions have been described [24, 25]. Respondents were asked their number of sexual partners in the previous 12 months for three partnership types (spouse/live-in; steady; casual), which we collapsed into two (Type I = spouse/live-in, Type II = steady/casual). For up to the four most recent partners per category, respondents were asked for date (by month) of the first and the most recent time they had sex with this person. These are the questions recommended by the UNAIDS reference group on Estimates, Modeling and Projection [26].

Point partnership distributions. A person's "point partnership count" is their number of ongoing partnerships at a given time. The survey did not ask whether partnerships were ongoing, so we could not calculate point partnership counts at interview date. However, partnership start- and end-dates allowed us to estimate counts for each month in the recent past, as recommended by

the UNAIDS reference group. We selected the period two months before interview as our cross-section, although this is only an approximation for multiple reasons; see Supplemental Digital Content (SDC). Table 2a shows the resulting point partnership distribution.

We used these data to construct an artificial network of 1,000 men and 1,000 women for model estimation; we selected a 1:1 sex ratio in the absence of precise data about the correct value for this population. An agent-based model of a closed, sex-balanced heterosexual population requires equal numbers of male and female partnerships. In our data, women report a point partnership mean of 0.62, compared to 0.70 for men, a difference of 0.08 partnerships. For the empirical population, this is likely due to it not being closed (i.e. men having more partnerships outside the target population), a greater proportion of men not fulfilling the study criterion of "ever sexually active", and/or the sex ratio not being 1:1 (with one report finding a nationwide sex ratio for ages 15-64 in Zimbabwe as low as 81 males per 100 females [27]). For our simulated population, which *is* sex-balanced and closed, we explore a base method for incorporating the data (by assuming the midpoint of sex-specific reports to be correct; see SDC) and a second method during the sensitivity analyses (see below).

Women also define a higher fraction of partnerships as Type I (54.9%) than do men (31.0%). We take the midpoints of the sex-specific reports as our estimates (see SDC). Implementing these assumptions led to the expected counts in Table 2b.

Durations: Questions on dates of first and last sex for partnerships two months before interview drove our initial estimates for average duration of Type I (45.3 months) and Type II (11.2

months) partnerships, assuming a 1/3-month duration for partnerships beginning and ending in the same month. Given the bias towards observing longer partnerships in a cross-sectional survey, we re-weighted partnerships according to their relative inclusion probability, equivalent to using the harmonic mean of observed durations.

Partnership model estimation. The above statistics are used to estimate a partnership network model, which then drives a stochastic network simulation that has these statistics as their expected values for partnership durations and sex- and relational type-specific concurrency. We use recent developments in exponential random graph modeling (ERGM) [28-35], with an extension for dynamic networks [36], to estimate network model parameters. In the cross-sectional ERGM framework, a population is represented by a network comprising *n* nodes (persons) and a matrix $\mathbf{Y}$ of pairwise partnerships (wherein $Y_{ij} = 1$ signifies a partnership between *i*,*j*; $Y_{ij} = 0$ signifies absence). One may also subset these matrices by nodal and/or dyadic attributes. The framework allows one to model probabilities of partnerships forming or dissolving as a function of individual nodal attributes (e.g. disease status), pairwise nodal attributes (e.g. age-based assortative mixing), partnership type, and relational configurations (e.g. each person's number of ongoing partnerships), and to estimate coefficients for these effects jointly from data. The general form is:

$$\text{logit}\left(Y_{ij} = y_{ij} \middle| n \text{ nodes}, \mathbf{x}_n, \mathbf{x}_d, \mathbf{y}_{ij}^c\right) = \theta^T \Delta g\left(\mathbf{y}_{ij}^c, \mathbf{x}_n, \mathbf{x}_d\right)$$

where y_{ij}^c is the complement of y_{ij} (the relational values for all actor pairs besides *i,j*); $\mathbf{x}_n$ and $\mathbf{x}_d$ are nodal and dyadic covariate matrices, respectively; $g(\mathbf{y}, \mathbf{x}_n, \mathbf{x}_d)$ is an arbitrary vector of statistics measured on $\mathbf{y}$; $\Delta g(\mathbf{y}_{ij}^c, \mathbf{x}_n, \mathbf{x}_d)$ is the change in those statistics when y_{ij} changes from 0 to 1; and θ is a vector of coefficients. Any set of statistics can be hypothesized to affect network probability and thus be included in $g(\mathbf{y}, \mathbf{x}_n, \mathbf{x}_d)$. The logit formulation highlights the partnership as unit of analysis. Although it resembles a simple logistic regression model, the dependence among observations (induced when **g** contains statistics such as point partnership counts) makes inference considerably more complex. For the discrete dynamic extension [36], the probability of $\mathbf{Y}$ at time *t* is a function of its cross-sectional statistics and of the y_{ij} values at *t-1*; this is accomplished using separate ERGM formulas for relational formation and dissolution.

We model a population with three nodal attributes (sex, HIV-status, time since infection) and one dyadic attribute (Type I or II). Statistics in **g** for our baseline formation model (each of whose index $(\mathbf{y}, \mathbf{x}_n, \mathbf{x}_d)$ is left off for readability) are:

$$\{e_{\square};\ u_{f,1,I};\ u_{f,1,\square};\ u_{m,1,I};\ u_{m,1,\square};\ u_{m,2,\square};\ v_{f,2,\square};\ v_{m,3,\square}\}$$

where:

- e_t = total partnerships of type *t*;
- $u_{s,n,t}$ = # of individuals of sex *s* in exactly *n* partnerships of type *t*

- $v_{s,n}$ = # of individuals of sex s in $\geq n$ partnerships of type t

and □ indicates both partnership types combined. Model constraints include $v_{f,2,I} = v_{m,3,I} = 0$.

The statistics in the **g** vector for our baseline dissolution model include:

$$\{e_{\square};\ e_I\}$$

which implies that all partnerships within a type have identical daily dissolution probabilities.

These statistics are constructed to capture the expected proportion of each sex with given Type I and Type II point partnerships counts, and expected durations of partnerships by type. Numerous model constraints mean that additional statistics beyond those explicitly included are fit automatically; see the SDC. Collectively, these provide a detailed specification of concurrency patterns.

Partnership network simulation. One strength of the ERGM framework is that network simulation is driven by the same model as estimation. We implement both processes using *statnet* (*http://www.statnetproject.org*), whose core algorithms employ Markov chain Monte Carlo (MCMC); fitting details used (e.g. burn-in, simulation sample) are available from the authors on request. The stochastic model estimated above drives dynamic network simulation, ensuring that the expected values of network statistics are preserved while the values themselves vary stochastically around these expectations at every cross-sectional slice.

Epidemic simulation. For each timestep, each serodiscordant couple's transmission probability is determined by the positive partner's time since infection. Given existing uncertainty about stage-specific transmission probabilities, we treat this as a sensitivity parameter. We consider four sets of transmission probabilities, outlined in Table A1: "Wawer" [1], "Pinkerton" [10], "Abu-Raddad" [11], and "Hollingsworth" [12]. The first three each contain three sub-models (High, Middle, Low) based on assumptions about coital frequency during the final five months of AIDS, on which the source papers lack data. We assume mean survival of 122 months after infection [12], and 40 years in the sexually active pool otherwise. A dying individual is immediately replaced by a new seronegative.

We initiate simulations with one male and one female infected. However, this implies that we are simulating an epidemic that starts with partnership rates observed in 2005. Since sexual behavior has likely changed over time with awareness of the epidemic's severity, it is important not to interpret our simulations' transitory dynamics as representing specific years. Our approach is best suited to projecting equilibrium dynamics for a given scenario. We run each simulation for 6000 months (500 years).

Alternative scenarios. We also considered three sensitivity or counterfactual scenarios, each using the Hollingsworth transmission estimates:

Male reports: here, rather than assuming the mid-point of male and female sexual behavior reports, we treat the difference as a function of female under-reporting. We repeat model

estimation and simulation using the slightly higher male-reported point partnership mean.

No-acute-peak: to explore the potential impact of an intervention eliminating the acute infection peak, we considered additional simulations in which Hollingsworth's chronic-stage transmission probability is also applied to the acute stage.

No-concurrency: to explore the potential role that concurrency plays in amplifying the impact of acute-stage transmission, and to consider potential impacts of a concurrency intervention, we estimated parameters for a model with identical point partnership mean and partnership durations as our data, but in which individuals' point partnership counts were constrained to 0 or 1. This preserves the total number of partnerships at any timepoint, and the total time spent in partnerships, from the baseline runs; effectively this reallocates observed concurrent partnerships to isolates to satisfy the serial monogamy constraint.

RESULTS

Point partnership mean was slightly higher for males than females (mean = 0.70 vs. 0.62), as previously discussed, as was standard deviation (0.779 vs. 0.577). Women were more likely to report exactly one partnership (55.6% vs. 43.4%); conversely, reports of >1 ongoing partnership were higher among men (11.6% vs. 2.8%), as were reports of no ongoing partnerships (45.0% vs. 41.6%). Overall, the point prevalence of concurrency was 7.3%.

Table A2 provides the coefficient estimates for our baseline model, and Figure A1 shows a

random cross-sectional draw from the network simulation. Statistics are consistent with the proposed probability model (e.g. 661 mean partnerships); the network is, in fact, quite sparse, and most members reside in cross-sectional relational components of 0-2 people.

Some models never generated a sustained run across 100 simulations. For those that did, Figure 1a shows mean HIV prevalence across five sustained runs. The Hollingsworth and Pinkerton-High models yield the only mean prevalence projections that stabilize above 5%. Hollingsworth's assumption of no coital acts during the last ten months of AIDS seems more realistic than assuming no coital reduction with the Pinkerton-High models; we thus adopt the Hollingsworth model as our base comparative model in subsequent analysis, conduct five additional simulations with it, and focus on it in more detail here (Figure 2). To see whether the proportion of transmissions stemming from each stage changes over time, we define the stages (for Figure 2a only) using Hollingsworth's classification (acute = 3 months; chronic = 100 months; AIDS = 19 months, including 10 without coital acts); we then plot mean fraction of transmissions by stage across the ten Hollingsworth runs against time. Although HIV prevalence stabilized very slowly (Figure 1a), the transmission proportions by stage stabilize very quickly. The latter are thus well described by the mean for each series, and we use this in subsequent analyses (excluding the first decade to eliminate any transient burn-in). Figure 2b shows a "cumulative transmission curve" for the age of all source infections in the ten base Hollingsworth runs. On average, about 22%, 49%, and 29% of transmissions stem from acute-, chronic- and AIDS-stage infections, respectively. There is very little variation across runs, given the large number of transmission events (mean = 4,736). Proportions of infection by stage are tabulated for these and all subsequent runs in Table A3.

Results from additional transmission probability scenarios are found in Figure 3 and Table A3. The Pinkerton and Wawer runs show a similar role for the acute stage as did Hollingsworth, while the Abu-Raddad runs suggest a lower role. All three models predict that more infections occur during AIDS stage (~40) than does the Hollingsworth model.

Using men's reports instead of the midpoint estimate of the point partnership mean increases this value by just 0.04 (from 0.66 to 0.70), and the fraction of women with concurrent partnerships rises 2.2 percentage points (from 2.8 to 5.0). This small change nearly doubles the equilibrium prevalence of the epidemic (Figure 2b). However, it does not change the fraction of transmissions by stage; the cumulative transmission curve is indistinguishable from baseline (not shown).

Assuming no acute infection peak (and using the midpoint behavioral estimates for concurrency), the epidemic appears to be unsustainable; it dies out long before 6,000 months in each of 100 simulations. Thus, even though the acute stage only contributes about 20% of infections (across the various transmission models), reducing transmissibility during acute infection down to chronic levels appears sufficient to send the epidemic into extinction.

The no-concurrency model also leads to epidemic extinction in all 100 runs. That is, if the partnerships reported in the survey were to occur sequentially instead of concurrently (but with the same mean degree of 0.66 partnerships in the cross section), and had the same durations and levels of coital frequency, the epidemic would not be sustainable.

If we look at the first 5% of the time series (300 months) for the runs that last this long, eliminating concurrency reduces the proportion of transmission stemming from acute infections (using Hollingsworth's three month definition) by half, down to 11.3%, before the epidemic dies out. Eliminating the acute infection peak reduces the proportion down to 0.0%.

DISCUSSION

Our models generally predict that in equilibrium, and with concurrency, ~20% of transmissions would stem from acute-stage infections in Zimbabwe, and ~30-45% from AIDS-stage infections. The former figure was not greatly affected by the exact stage-specific transmission probabilities assumed among published estimates, with the exception of the Abu-Raddad estimates. By comparison, Hollingsworth et al. [12] estimate 31% under random mixing, and Abu-Raddad and Longini [11] estimate ~5-13% at equilibrium (as seen in their Figures 1b and 2b). Pinkerton [10] estimates that 89.1% of the infections occurring during the first 20 months occur during the acute phase; for the same period our model shows 76% using Pinkerton's parameters and 75% using Hollingsworth's. All of these differences are likely due to the dramatically different ways in which we modeled behavior. Unlike all previous models, our approach incorporates the observed timing and sequence of an individual's multiple partnerships (whether concurrent or sequential), which should affect the probability of acquiring infection from one person and transmitting to another during a narrow time period. Differences may also result from real differences in behavior among different sub-Saharan African populations, although comparable data and models on relational timing in other settings would be needed to answer this question for sure.

Although only a small fraction of infections directly stemmed from acute index cases, we find that eliminating this peak would drive the epidemic below the reproductive threshold and into extinction, at least for this population. The same is true for concurrent partnerships; eliminating concurrency, while keeping cumulative partner numbers, durations, and coital frequencies the same, leads to epidemic extinction. In part, this is because eliminating concurrency reduces the impact of acute infection, cutting the proportion of infections from this stage in half before extinction occurs. These observations suggest that both behavioral and biomedical interventions reducing the impact of the acute stage could have enormous population-level impacts.

The Abu-Raddad numbers seemed least able to sustain a realistic epidemic, and the Hollingsworth numbers the most. The epidemic created by the Hollingsworth numbers, based on the average of male and female reports of ongoing partnerships, was still small. However, using the average number of ongoing partnerships reported by males increased the epidemic substantially. The difference in behavior was not large: an increase in 0.04 partners per person in the cross-section, and a 2.2 percentage point increase in the number of women with concurrent partners. This small change in behavior led to a 100% increase in equilibrium prevalence. This suggests that transmission in this population is very close to a behavioral "tipping point" wherein small changes in partnership rates, particularly concurrent partnerships, can have large effects on network connectivity and HIV prevalence.

This has implications both for prevention and for understanding the epidemiology of HIV in Africa generally. In order to match observed epidemic trajectories, previous modeling studies

that have not dealt explicitly with concurrency have had to posit the existence of significant fractions of the population with very high numbers of partners every year sustained over many years. Examples include 6.7% of the population in Yaoundé having an average of 2,870 partnerships per person, each at least one week long, over 35 sexually-active years [11], and 10% of the population in sub-Saharan Africa generally having an average of 723 partnerships per person over 40 years, each with 25 sex acts or more [37]. These behavioral patterns are not based on observed data, but they are apparently necessary, in the absence of concurrency, to generate realistic epidemics. The strong disconnect between these assumptions and empirical data has been criticized [38], and has been used to argue for the importance of non-sexual transmission routes in Africa [39]. The findings here suggest a different interpretation. Our work adds to earlier studies [40] showing that far smaller numbers of partners are necessary to generate substantial HIV epidemics when concurrency is explicitly measured and modeled. In countries where concurrency for women is stigmatized, it is likely that this leads to some level of social desirability bias and underreporting. Our findings also suggest that correcting for even a tiny amount of underreporting among women can yield large increases in predicted equilibrium prevalence, and generate realistic epidemics.

Our model included numerous simplifying assumptions that must be remembered when making inferences. Our data source, and thus our model, excluded people who had never had sex, as well as commercial sexual contacts. Each of these should affect our prevalence estimates, albeit in opposite directions. We did not consider additional network structuring induced by geographic mixing, nor mixing by age or other exogenous demographic or social variables. We did not consider variation in circumcision status or co-infection with STDs or other infections.

Our estimates for concurrency are subject to several sources of measurement error. Some may lead to overestimates (e.g., dates reported by month only), others to underestimates (e.g. low coital frequency, social desirability bias). Our estimates of stage-specific transmission probabilities and coital frequency come from a single study, and are also subject to measurement error. There remains uncertainty in the duration and magnitude of the acute infection window, and in coital frequency during late-stage AIDS. For many of these limitations, the sensitivity analyses we conduct on transmission probabilities and behavioral parameters can give some intuition regarding the impact of these factors on model outcomes.

Finally, all of the behavioral parameters – especially those used for network simulation – are population-specific. This is true for all modeling studies, and the constraint this places on inference is often underestimated. Given the dramatic impacts of small changes in behavioral assumptions on equilibrium prevalence that we found, it is clear that estimates of the stage-specific fraction of infections, from any modeling study, must be limited to the population from which the behavioral data are drawn.

Despite these limitations, we were able to model sexual behavior and patterns of partnership timing and overlap much more accurately than previous studies, and this clearly matters for understanding transmission during acute infection. A short acute infection window will always have a strong interaction with rapid serial partner acquisition and concurrency in sexual transmission dynamics. These patterns of relational timing vary widely among populations in which HIV is spread, and it is therefore important to model them with some fidelity. The new modeling tools that we employ here can do that, using simple egocentrically sampled network

data. The insights these new methods provide are critical for "knowing your epidemic," and for developing population-specific prevention packages.

ACKNOWLEDGMENTS

The authors would like to thank the study participants as well as Mark Handcock, David Hunter, Pavel Krivitsky, Carter Butts, and the entire *statnet* development team.

SMG and SC were supported in part by the Puget Sound Partners for Global Health (Research and Technology Project Award 26145). Data collection was supported by NIH (U10MH061544 and RZIAA014802). The models and software used in this study were developed with support from NIH (R01 HD041877 and R01 DA12831). SC was partially supported by NIH (K99 HD057553 and P30 AI27757).

SMG was responsible for conceiving of the modeling project, conducting the majority of the modeling, and primary preparation of the manuscript. SC conducted secondary data analysis and aided in the modeling work and manuscript preparation. DK, DEM, and AG designed the survey instrument, oversaw the data collection, conducted the primary data analysis, and provided assistance in manuscript preparation. MM contributed to project conception, data analysis, and manuscript preparation.

TABLE 1: Previous studies on proportion of HIV infections by stage of index case

Author (year)	**Population**	**Method**	**% infections from acute stage**	**Duration of acute stage**	**Comments**
Pinkerton (2008) [10]	Heterosexual transmission, Uganda	computer-based model, likelihood estimate (Bernoulli transmission model)	--	49 days	Pinkerton does not estimate proportion of all infections stemming from acute stage, but does state that his model predicts that out of the infections occurring in the first 20 months, 89.1% are due to acute stage index cases. This number is based on the 12 transmission events in the "incident index partner" couples from Wawer (2005)
Hollingsworth et al. (2008) [12]	Heterosexual transmission, Uganda	Parametric model-based likelihood estimation	9% , 31%	2.9 months	9% under assumption of serial monogamy; 31% under assumption of random mixing
Abu-Raddad and Longini (2008) [11]	Heterosexual transmission, sub-Saharan Africa	Deterministic compartmental mathematical model	13%, 18%	2.5 months	These estimates are the *cumulative* proportions by approximate equilibrium (years 2050 and 2080) for Kisumu and Yaoundé, respectively, and thus include the early epidemic as well. The proportion by stage per unit time at endemic equilibrium appears to be ~8% and ~15%, respectively (Figures 1b and 2b).
Hayes & White (2006) [14]	Heterosexual transmission, Uganda	Details unknown [not described in letter]	23%, 41%	5 months	41% if each seroconverter has one subsequent seronegative partner; 23% if all of each seroconverter's subsequent partners are susceptible (such as a sex worker in a low-prevalence setting).
Cohen & Pilcher (2005) [13]	Heterosexual transmission, Uganda	Review of Wawer 2005 data	~50%	5 months	Interpreting this number as a population-based estimate may be inaccurate. See, for instance, Pinkerton et al.'s discussion ()

Pinkerton (2007) [15]	US population	A "simple" multiplicative mathematical model	8.6%	49 days	In a variety of sensitivity analysis, estimates ranged from 2.5 - 17.3%.
Prabhu (2009) [16]	US population	Same as Pinkerton (2007)	11.4%	49 days	Prabhu et al updated Pinkerton (2007) with new HIV incidence data in the US
Xiridou et al (2004) [18]	MSM in Amsterdam	set of differential equations	11%	3 months	35% of HIV transmissions from casual partners and 6% of transmissions from steady partners
Rapatski et al. (2005) [17]	MSM in San Francisco	Piece-meal model: likelihood estimates, then deterministic, compartmental model with stage of disease	1.3%	6 months	This modeling work suggested that the late-stage, may account for the majority of disease transmission (97%) in settings with established epidemics; however, these findings have been disputed in the literature [41].
Brenner et al. (2007) [21]	Majority MSM, Urban Quebec	Population-based phylogenetic approach	49.4%	N/A	
Pao et al. (2005) [19]	Majority MSM, United Kingdom	Population-based phylogenetic approach	34%	N/A	
Yerly et al. (2001) [20]	Swiss HIV cohort study (42% MSM, 13% IDU)	Population-based phylogenetic approach	30%	N/A	

TABLE 2: Cross-sectional network statistics

a) POINT PARTNERSHIPS BY TYPE

FEMALES

# of Type I partnerships \ # of Type II partnerships	0	1	2	3	4	5	6	
0	843	455	32	6	2	0	1	1,339
1	673	12	1	0	0	0	0	686
2	1	0	1	0	0	0	0	2
	1,517	467	34	6	2	0	1	2,027

MALES

# of Type I partnerships \ # of Type II partnerships	0	1	2	3	4	5	6	
0	978	576	128	25	11	0	0	1,718
1	366	49	16	4	1	0	0	436
2	11	4	0	0	0	0	0	15
3	1	0	0	0	0	0	0	1
4	1	0	0	0	0	0	0	1
	1,357	629	144	29	12	0	0	2,171

TOTAL POINT PARTNERSHIPS		0	1	2	3	4	5	6	Mean	Std. dev.
	F	843	1128	45	7	3	0	1	0.620	0.577
	M	978	942	188	46	16	1	0	0.702	0.779

b) POINT PARTNERSHIPS BY TYPE

FEMALES

# of Type I partnerships \ # of Type II partnerships	0	1	2	3+	
0	0.399	0.289	0.034	0.005	0.727
1	0.262	0.010	0.000	0.001	0.273
	0.661	0.299	0.034	0.006	1.000

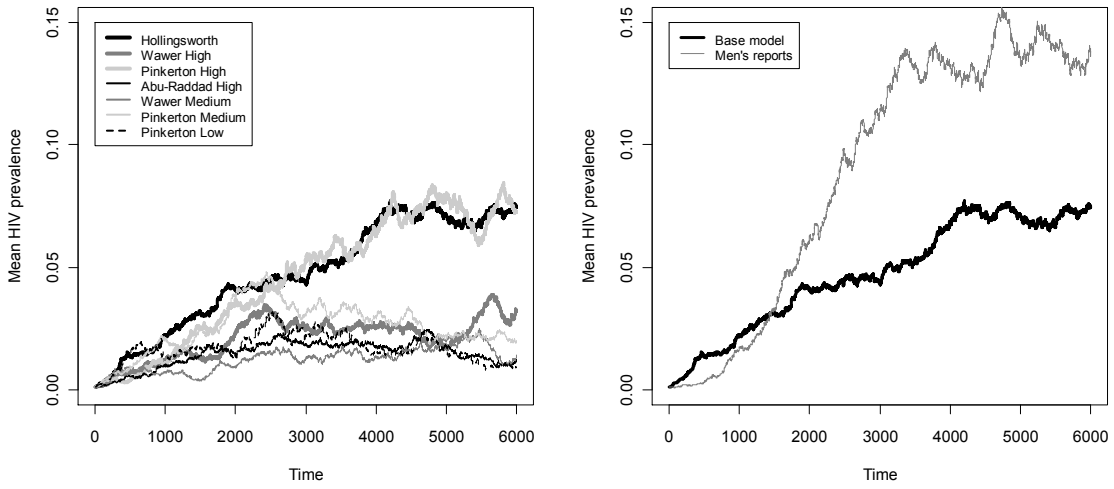
MALES

# of Type I partnerships \ # of Type II partnerships	0	1	2	3+	
0	0.483	0.208	0.039	0.011	0.741
1	0.200	0.033	0.009	0.003	0.245
2	0.009	0.004	0.001	0.000	0.014
	0.692	0.245	0.049	0.014	1.000

TOTAL POINT PARTNERSHIPS		0	1	2	3+	Mean	Std. dev.
	F	0.399	0.551	0.044	0.006	0.661	0.615
	M	0.483	0.408	0.072	0.023	0.661	0.772

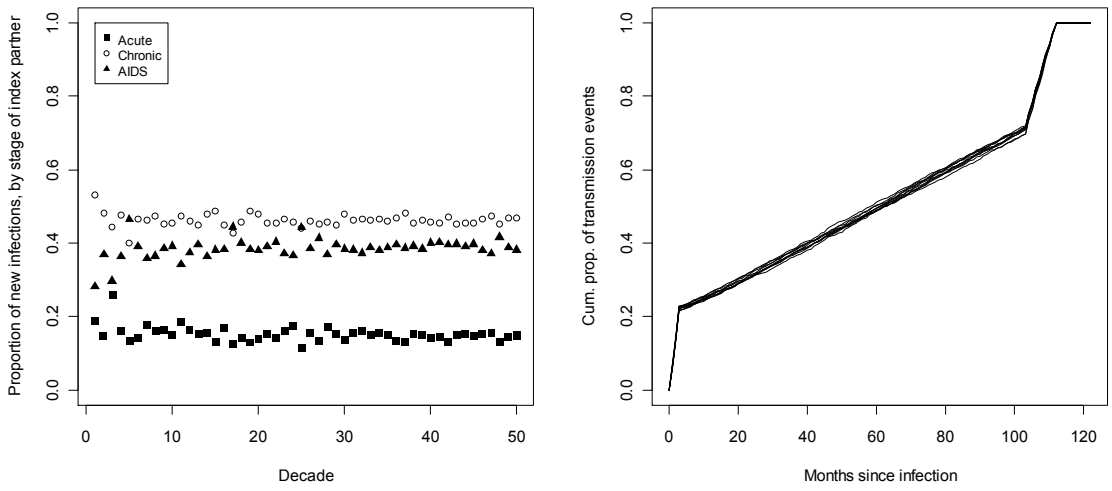
a) observed in the Zimbabwe survey data. Figures represent counts.
b) reconciled for model estimation. Figures represent proportions of the sex-specific population.

FIGURE 1: HIV prevalence



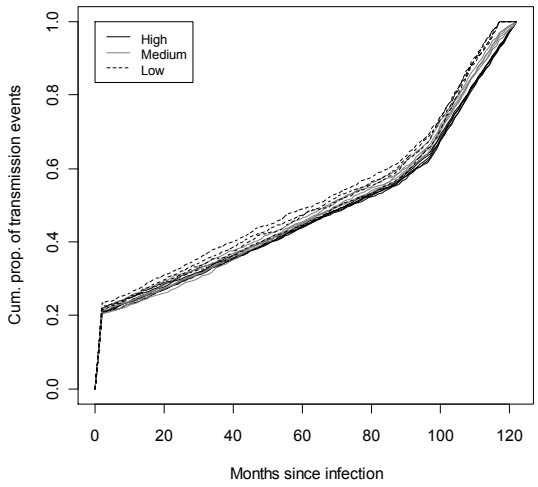
a) Transmission models. All Wawer-Low, Abu-Raddad-Medium, and Abu-Raddad-Low runs die out.
b) Behavioral models. All no-acute-peak and no-concurrency runs die out quickly.

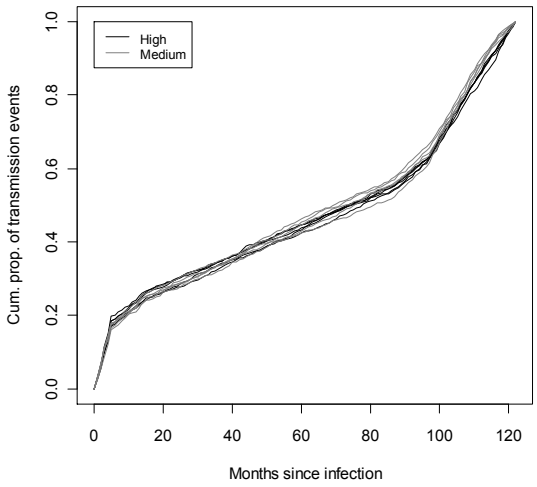
FIGURE 2: Hollingsworth model

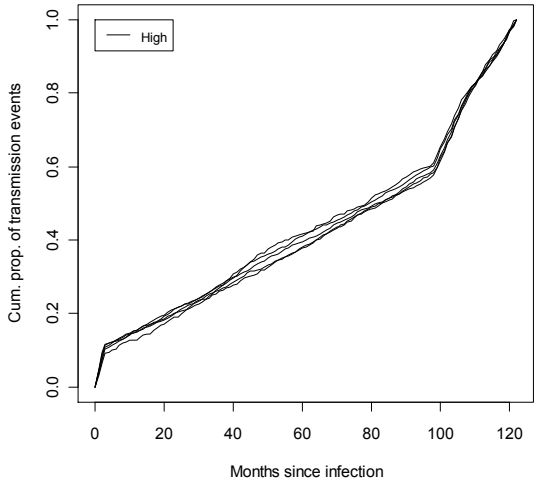


a) Proportion of new infections by stage over time, averaged across all ten Hollingsworth base model runs
b) Cumulative transmission curve, i.e. the cumulative proportion of transmissions stemming from infections no older than a given age. Each line represents one of ten Hollingsworth base model runs. The first 10 years of simulation "burn-in" are excluded. The horizontal line at the end of the chart is due to the assumption of no sex during the last ten months of infection for this model.

FIGURE 3: Other transmission models







(a) Pinkerton
(b) Wawer
(c) Abu-Raddad

Each line represents one simulation, with five simulations per scenario. The Wawer-Low, Abu-Raddad-Medium, and Abu-Raddad-Low runs consistently die out and are not shown.

1. Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li XB, Laeyendecker O*, et al.* **Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda**. *J Infect Dis* 2005,191:1403-1409.
2. Coombs RW, Speck CE, Hughes JP, Lee W, Sampoleo R, Ross SO*, et al.* **Association between culturable human immunodeficiency virus type 1 (HIV-1) in semen and HIV-1 RNA levels in semen and blood: Evidence for compartmentalization of HIV-1 between semen and blood**. *J Infect Dis* 1998,177:320-330.
3. Leynaert B, Downs AM, de Vincenzi I. **Heterosexual transmission of human immunodeficiency virus - Variability of infectivity throughout the course of infection**. *American Journal Of Epidemiology* 1998,148:88-96.
4. Little S, McLean A, Spina C, Richman D, Havlir D. **Viral dynamics of acute HIV-1 infection**. *J Exp Med* 1999,190:841-850.
5. Quinn TC, Wawer MJ, Sewankambo N, Serwadda D, Li CJ, Wabwire-Mangen F*, et al.* **Viral load and heterosexual transmission of human immunodeficiency virus type 1** *N Engl J Med* 2000,343:921-929.
6. Jacquez JA, Koopman J, Simon CP, Longini IMJ. **Role of the primary infection in epidemics of HIV-infection in gay cohorts**. *J Acquired Immune Defic Syndromes Hum Retrovirol* 1994,7:1169-1184.
7. Koopman JS, Jacquez JA, Welch GW, Simon CP, Foxman B, Pollock SM*, et al.* **The role of early HIV infection in the spread of HIV through populations**. *J Acquired Immune Defic Syndromes Hum Retrovirol* 1997,14:249-258.
8. Longini IMJ, Clark WS, Haber M, Hosburgh RJ. **The stages of HIV infection: waiting times and infection transmission probabilities**. *Lecture Notes in Biomathematics* 1989,83:111-137.
9. Pilcher CD, Tien HC, Eron JJJ, Vernazza PL, Leu S-Y, Stewart PW*, et al.* **Brief but efficient: Acute HIV infection and the sexual transmission of HIV** *The Journal of Infectious Diseases* 2004,189:1785-1792.
10. Pinkerton SD. **Probability of HIV transmission during acute infection in Rakai, Uganda**. *AIDS and Behavior* 2008,12:677-684.
11. Abu-Raddad LJ, Longini IA. **No HIV stage is dominant in driving the HIV epidemic in sub-Saharan Africa**. *AIDS* 2008,22:1055-1061.
12. Hollingsworth TD, Anderson RM, Fraser C. **HIV-1 transmission, by stage of infection**. 2008.
13. Cohen MS, Pilcher CD. **Amplified HIV transmssion and new approaches to HIV Prevention**. *J Infect Dis* 2005,191:1391-1393.
14. Hayes R, White R. **Amplified HIV transmissoin during early-stage inefction [letter]**. *J Infect Dis* 2005,193:604-605.
15. Pinkerton SD. **How many sexually-acquired HIV infections in the USA are due to acute-phase HIV transmission?** *AIDS* 2007,21:1625-1629.
16. Prabhu VS, Hutchinson AB, Farnham PG, Sansom SL. **Sexually acquired HIV infections in the United States due to acute-phase HIV transmission: an update**. *AIDS* 2009,23:1792-1794.

17. Rapatski BL, Suppe F, Yorke JA. **HIV Epidemics Driven by Late Disease Stage Transmission**. *J Acquir Immune Defic Syndr* 2005,38:241-253.
18. Xiridou M, Geskus R, de Wit J, Coutinho R, Kretzschmar M. **Primary HIV infection as source of HIV transmission within steady and casual partnerships among homosexual men**. *AIDS* 2004,18:1311-1320.
19. Pao D, Fisher M, Hue S, Dean G, Murphy G, Cane PA*, et al.* **Transmission of HIV-1 during primary infection: relationship to sexual risk and sexually transmitted infections**. *AIDS* 2005,19:85-90.
20. Yerly S, Vora S, Rizzardi P, Chave JP, Vernazza PL, Flepp M*, et al.* **Acute HIV infection: impact on the spread of HIV and transmission of drug resistance**. *AIDS* 2001,15:2287-2292.
21. Brenner BG, Roger M, Routy JP, Moisi D, Ntemgwa M, Matte C*, et al.* **High rates of forward transmission events after acute/early HIV-1 infection**. 2007.
22. Morris M, Goodreau SM, Moody J. **Sexual Networks, Concurrency, and STD/HIV**. In: *Sexually Transmitted Diseases*. Edited by Holmes KK. New York: McGraw-Hill 2007:109 - 126.
23. Morris M, Kurth AE, Hamilton DT, Moody J, Wakefield S, The Network Modeling Group. **Concurrent partnerships and HIV prevalence disparities by race: Linking science and public health practice**. *American Journal of Public Health* 2009,99:1023-1031.
24. NIMH Collaborative HIV/STD Prevention Trial Group. **Methodological overview of a five-country community-level HIV/sexually transmitted disease prevention trial**. *AIDS* 2007.
25. Kasprzyk D, Montaño DE. **Application of an integrated behavioral model to understand HIV prevention behavior of high-risk men in rural Zimbabwe**. In: *Prediction and change of health behavior: applying the reasoned action approach*. Edited by Ajzen I, Hornik R. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.; 2007.
26. UNAIDS Reference Group on Estimates Modelling and Projections. **Consultation on Concurrent Sexual Partnerships: Recommendations from a meeting of the UNAIDS Reference Group on Estimates, Modelling and Projections held in Nairobi, Kenya, April 20-21st 2009** In. http://www.epidem.org/Publications/Concurrency%20meeting%20recommendations_Final.pdf; 2009.
27. Central Intelligence Agency. **CIA World Factbook - Zimbabwe**. In; 2009.
28. Frank O, Strauss D. **Markov graphs**. *Journal of the American Statistical Association* 1986,81:832-842.
29. Strauss D, Ikeda M. **Pseudolikelihood estimation for social networks**. *J Am Stat Soc* 1990,85:204-212.
30. Wasserman S, Pattison P. **Logit models and logistic regressions for social networks: I. an introduction to Markov graphs and p***. *Psychometrika* 1996,60:401-425.
31. Handcock MS. **Statistical models for social networks: degeneracy and inference**. In: *Dynamic Social Network Modeling and Analysis : workshop summary and papers*. Edited by Breiger RL, Carley KM, Pattison P. Washington, D.C.: National Academies Press; 2003:229-240.
32. Snijders TAB, Pattison PE, Robins GL, Handcock MS. **New specifications for exponential random graph models**. *Sociological Methodology* 2006,36:99-153.

33. Robins G, Morris M. **Advances in exponential random graph (p*) models**. *Social Networks* 2007,29:169-172.
34. Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M. **Statnet: software tools for the representation, visualization, analysis and simulation of social network data**. *Journal of Statistical Software* 2008,24:1-11.
35. Hunter DR, Goodreau SM, Handcock MS. **Goodness of fit of social network models**. *Journal of the American Statistical Association* 2008,103:248-258.
36. Krivitsky PN. **Statistical Models for Social Network Data and Processes**. Seattle: University of Washington; 2009:1-180.
37. 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:e2212.
38. Deuchert E, Brody S. **Plausible and implausible parameters for mathematical modeling of nominal heterosexual HIV transmission**. *Ann Epidemiol* 2007,17:237-244.
39. Gisselquist D, Rothenberg R, Potterat J, Drucker E. **HIV infections in sub-Saharan Africa not explained by sexual or vertical transmission**. *Int J STD AIDS* 2002,13:657-666.
40. Morris M, Kretzschmar M. **Concurrent partnerships and the spread of HIV**. *AIDS* 1997,11:641-683.
41. Koopman JS, Simon CP. **Response to Rapatski BL, Suppe F, Yorke JA. HIV epidemics driven by late disease stage transmission**. *Jaids-Journal of Acquired Immune Deficiency Syndromes* 2006,41:677-677.