Epidemics 2 (2010) 155–164

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Epidemiological bridging by injection drug use drives an early HIV epidemic

Erik Volz [a,*], Simon D.W. Frost [b], Richard Rothenberg [c], Lauren Ancel Meyers [d,e]

[a] Department of Epidemiology, University of Michigan – Ann Arbor, USA
[b] Department of Veterinary Medicine, University of Cambridge, Cambridge, UK
[c] Institute of Public Health, Georgia State University, Atlanta, GA, USA
[d] Department of Integrative Biology, University of Texas – Austin, USA
[e] Santa Fe Institute, Santa Fe, NM, USA

ARTICLE INFO

Article history:
Received 6 March 2010
Revised 12 May 2010
Accepted 17 June 2010

Keywords:
HIV
Networks
Serosorting
Concurrency

ABSTRACT

The risk of acquiring sexually transmitted infections (STIs) depends on individual behavior and the network of risky partnerships in which an individual participates. STI epidemics often spread rapidly and primarily among individuals central to transmission networks; and thus they often defy the mass-action principle since incidence is not proportional to the infectious fraction of the population. Here, we estimate the contact network structure for an Atlanta, Georgia community with heterogeneous sexual and drug-related risk behaviors and build a detailed transmission model for HIV through this population. We show that accurate estimation of epidemic incidence requires careful measurement and inclusion of diverse factors including concurrency (having multiple partners), the duration of partnerships, serosorting (preference for partners with matching disease state), and heterogeneity in the number and kinds of partners. In the focal population, we find that injection drug users (IDUs) do not directly cause many secondary infections; yet they bridge the heterosexual and men-who-have-sex-with-men (MSM) populations and are thereby indirectly responsible for extensive transmission.

The role of bridging populations has been noted in both the theoretical and empirical (Anderson and May, 1991; Keeling and Rohani, 2008) study of many infectious diseases. For example, several recent studies (Kral et al., 2001; Pisani et al., 2003; Howard and Latkin, 2006; Lindenburg et al., 2006; Kretzschmar and Wiessing, 2008) on the role of injection drug use in the HIV pandemic have found that injection drug users (IDUs) constitute a high-risk core group which facilitate transmission to lower risk populations with sexual risk factors. Although the impact of injection drug use has largely been diminished by public health interventions and herd immunity in mature HIV epidemics in North America and Europe (Lansky and Drake, 2009), it still accounts for a large fraction of transmissions in emerging HIV epidemics throughout the developing world (Grassly et al., 2003; Abramovitz et al., 2009). In the case of sexually transmitted diseases, the direct quantification of the role of bridge populations is complicated by many factors, including the complex network structure of sexual contacts that mediate transmission (Saidel et al., 2003; Wiessing and Kretzschmar, 2003) and survey designs which typically yield biased cross-sectional snapshots of a dynamic network structure (Salganik and Heckathorn, 2004; Volz and Heckathorn, 2008; Rothenberg et al., 2000).

Here, we have sought to overcome these problems and quantify the epidemiological roles played by various risky behaviors using data from a social network study that included 226 individuals at high risk for HIV infection in Atlanta, Georgia (Rothenberg et al., 2000). These data were used to parameterize a realistic network-based mathematical model of HIV transmission dynamics that accounts for extended-duration partnerships with different risk factors and thus allowed us to investigate the epidemiological significance of each partnership type. The model also incorporates the natural history of HIV infection and can accommodate a diverse range of network topologies including arbitrary degree distributions, variable duration of partnerships, and the tendency of HIV positive individuals to positively assort with other seropositives. Subjects were recruited into this study over three years using a chain-referral survey method known to produce biased samples (Salganik and Heckathorn, 2004; Volz and Heckathorn, 2008; Rothenberg et al., 2000). We corrected sample bias using sociometric variables to derive inverse probability weights, and then estimated the structure of the sexual and drug-use contact network, including the variable numbers of partners with different sexual and drug-use risk factors, variable durations of partnerships, and variable contact rates per partnership.

Our analysis suggests that injection drug use can play an important and changing epidemiological role in this community. Although HIV spreads more rapidly among IDUs than among other risk groups in the community, transmissions through injection drug use never account for the majority of transmissions and rapidly fall off in later stages of the epidemic. Despite this, injection drug use greatly exacerbates the ultimate

* Corresponding author.
E-mail address: erikvolz@umich.edu (E. Volz).

1755-4365/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.epidem.2010.06.003

extent of the epidemic by bridging the community of men-who-have-sex-with-men (MSM) and the predominately heterosexual community (HET). This essential epidemiological role of IDU's, however, is only apparent through analyses that properly accounts for network structure.

Methods

Our data, which are described in great detail in Rothenberg et al. (2000), were collected using a chain-link sample design. A partial representation of the network of partnerships revealed by the survey is illustrated in Fig. 1. Seeds for each of the six constructed chains were chosen as representative of the neighborhood based on ethnographic evaluation. These persons were over 18 years of age, male or female, and were involved in activities that placed them at risk for HIV acquisition or transmission. The study attempted to interview all participants at least twice, and succeeded with 80% of respondents. A third interview was conducted with 50% of respondents. Longitudinal sampling allowed for the identification of concurrent partnerships as described below.

The sample composition is not generalizable to the population-at-large, whose risk behavior is likely to be much less intense. We hypothesize that the sample is representative of inner city at-risk groups in Atlanta.

Respondents in the survey were able to report multiple concurrent partnerships which sometimes featured multiple risk factors. In each partnership, we recorded the primary risk factor, which could be anal intercourse (AI), vaginal intercourse (VI), or sharing needles when injecting drugs (NS). The primary risk factor (PRF) for each partnership was determined by following a hierarchy of risk factors in the order of NS, AI, and VI. These simplifications were necessary as the full spectrum of risk behaviors within a single partnership can be very complex and difficult to operationalize in a mathematical model. The PRF was considered to be needle sharing if any sharing of injection paraphernalia was reported and regardless of sexual behavior. Sexual risk behavior was reported in only six of 53 needle-sharing partnerships, so there was not much overlap of risk behavior in this sample. AI was considered the PRF if there was no needle sharing and regardless of whether there was VI. And VI was the PRF otherwise. Sexual partnerships in which perfect condom use was reported were excluded from the analysis. And drug injection partnerships were excluded if no needle sharing was reported. We then estimated two quantities:

- the mean duration of each partnership type (the interval from first to final contact with a given partner), and
- the mean contact rate (contacts per day) of each partnership type.

Analyses were based on both the point estimate of these quantities and a sensitivity analysis using estimated standard errors.

We estimated the structure of the contact network using sampling theory developed for Respondent Driven Sampling (Salganik and Heckathorn, 2004; Volz and Heckathorn, 2008), a chain-referral sampling method. Survey weights were calculated that account for sample bias due to over-recruitment of individuals with high sociometric degree. For the purpose of deriving sample weights, we defined sociometric degree to be the number of partnerships reported during an interview, regardless of whether those partnerships are used for subsequent analysis. This includes partnerships of a social nature, which lack epidemiological significance, but could nevertheless influence the probability of being included in the survey.

Weights are inversely proportional to the probability that a unit is included in the sample. The sample inclusion probability is assumed to be proportional to sociometric degree of a unit, d_i, since more partnerships create more avenues for recruitment into the study. Therefore weights are proportional to $1/d_i$. Justification for this idea can be found in Salganik and Heckathorn (2004) and Volz and Heckathorn (2008). We further weighed each unit by a factor that corrects for homophilous recruitment tendencies, called a *recruitment weight*, which we denote ω. This accounts for the possibility that if a certain unit is sampled, similar units will subsequently be sampled and that moreover, such units may generate more recruitments on average. These patterns can, for example, skew a sample towards units similar to the *seeds* that begin the recruitment trees. Unfortunately, ω does not have a closed form, but our approach does not deviate significantly from the one presented by Volz and Heckathorn (2008). The total weight for unit i is $w_i = \omega_i / d_i$.

Weighted means are used as estimates for model parameters such as concurrency and partnership durations. For example, let $\{w_i\}$ be the set of derived weights for each sample unit i, and $\{r_i\}$ be the set of average durations of partnerships for each sample unit. Then, we would estimate the average duration of partnerships as

$$1/\hat{\rho} = \sum_i w_i \times r_i / \sum_i w_i. \quad (1)$$

In Section 1.1, we describe how weights are incorporated into the degree distributions used for modeling. Standard deviations were estimated using the variance estimator presented (Volz and Heckathorn, 2008), which is a refinement of the Hansen and Hurwitz (1943) variance estimator for chain-referral samples.

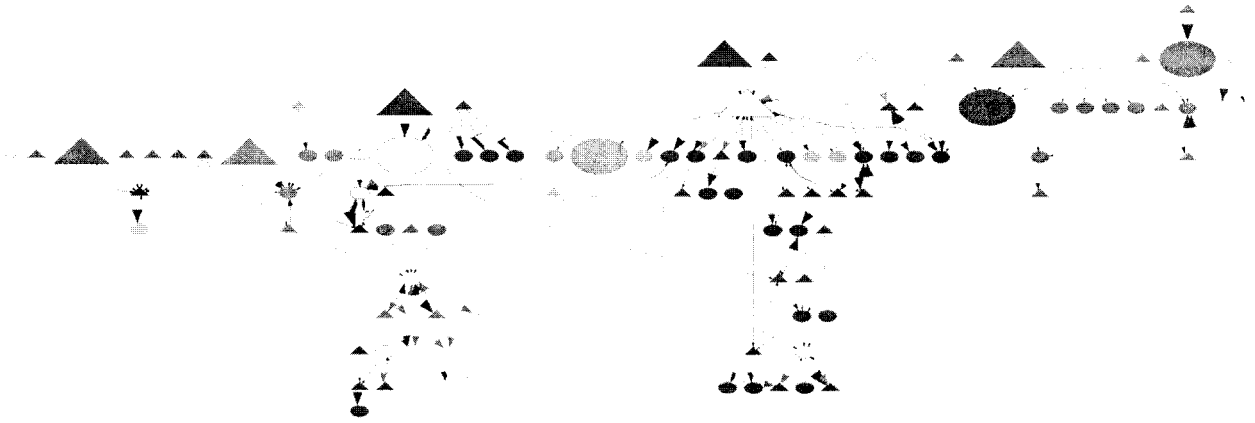

Fig. 1. The network of risky partnerships reported by participants in the Atlanta Urban Networks study. Triangles are males, and ovals are females. Large nodes represent HIV positive individuals. Red links represent partnerships such that the primary risk factor is sharing needles when injecting drugs (NS). Green and blue links respectively represent anal intercourse (AI) and vaginal intercourse (VI). The direction of arrows indicates that a study participant (the origin of the arrow) reported contact with the target of the arrow. The color of nodes represents the order in which subjects were recruited into the study (dark shades were recruited earlier and light shades later). Individuals that report any injection drug use have white interiors.

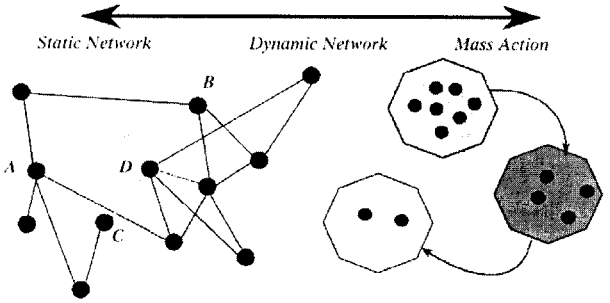

Fig. 2. A schematic of the dynamic network model. Edges are selected at random from the network at a constant rate and exchanged. The two dashed orange edges are selected and replaced by the two solid orange edges. As the rate of this process, ρ, is increased, the system reduces to mass-action dynamics with no static network structure.

Network model

A large body of theory has been developed for static random networks (Newman et al., 2001; Callaway et al., 2001; Catanzaro et al., 2005) with arbitrary degree distributions and the spread of epidemics on such networks (Volz and Meyers, 2007; Volz and Meyers, 2009; Meyers, 2003; Meyers et al., 2006). We will focus on random networks generated by the *configuration model* (*CM*) (Molloy and Reed, 1998) with degree distribution $\{p_k\}$. Since we will often need to refer to the two nodes that constitute an edge, we will use the terminology *ego* and *alter* to refer to the nodes under consideration. $K(ego)$ is a random variable describing the degree of node *ego*. $K(ego)$ can be interpreted as the number of simultaneous or concurrent contacts experienced by *ego*, and we will call it the *concurrent degree*.

In the limit of large network size, our networks have the property that given an edge from node *ego* to node x, the probability that $x = alter$ is proportional to the degree of node alter (k_{alter}). Below, we use $\{ego, alter\}$ to denote the edge connecting nodes *ego* and *alter* to each other and $(ego, alter)$ to denote a directed half-edge from *ego* to *alter*. Half-edges are potential conduits for transmission, so that if the half-edge $(ego, alter)$ exists, *ego* may potentially transmit to *alter*.

Our model considers only undirected networks in which every edge $\{ego, alter\}$ consists of two half-edges $(ego, alter)$ and $(alter, ego)$.

In Volz and Meyers, (2007), we introduced a class of dynamic network that is a generalization of CM networks. In these networks, edges are constantly rearranged by a *neighbor-exchange* process. A neighbor exchange can be represented by the following pseudo-chemical equation and is represented in Fig. 2:

$$\{ego, alter\} + \{ego', alter'\} \longrightarrow \{ego, alter'\} + \{ego', alter\}. \quad (2)$$

A given half-edge, $(ego, alter)$, will undergo a neighbor-exchange at a constant rate ρ. At a rate ρ, the half-edge $(ego, alter)$ will be transformed to $(ego, alter')$ such that $alter'$ is a node selected with probability proportional to the degree of $alter'$. This process preserves the concurrent degree sequence of the network.

We have extended the model so that during a neighbor-exchange, individuals can preferentially attach to others with matching serostatus, for example, a diagnosed HIV positive individual may seek new partners who are also HIV positive. This phenomenon, which has been dubbed *serosorting* (Cassels et al., 2009; Butler and Smith, 2007), is thought to play a large role in modulating HIV incidence in the modern epidemic.

To model this, partnerships between serodiscordant couples terminate at a constant rate μ and are then replaced with seroconcordant partnerships. A pseudo-chemical equation illustrating this process is the following:

$$2\{S, I\} \longrightarrow \{S, S\} + \{I, I\}. \quad (3)$$

There was insufficient information within the behavioral survey data to estimate serosorting rates directly. Therefore we indirectly inferred serosorting rates by choosing the value that is most consistent with the observed patterns of connectivity between susceptibles and infecteds and the sample HIV prevalence. The methods used are described in detail in Appendix B.

Nodes in the network progress through four states: susceptible, acute or early HIV infection, chronic or asymptomatic HIV infection, and removal or death. Infected nodes transmit to their concurrent partners in the network at a constant rate.

The transmission rates for acute and chronic infecteds to their susceptible partners are given in Table 1, along with other transmission probabilities that are based on published estimates (Baeten and Overbaugh, 2003; Leynaert et al., 1998; Kaplan and O'Keefe, 1993; White et al., 2007).

The model assumes that the acute or early HIV infection stage lasts on average 60 days (Vanhems et al., 2000), and the chronic infectious period lasts on average 10 years (Morgan et al., 2002).

We model transmission via three distinct pathways: the sharing of needles (NS) during injection drug use, anal intercourse (AI), and vaginal intercourse (VI).

Table 1
Model parameters and parameter estimates.

Parameter	Mean	Std. Dev.	Symbol
Avg. duration of acute infection (Vanhems et al., 2000)	60 days	–	$1/\gamma_I$
Avg. duration of chronic infection (Morgan et al., 2002)	10 years	–	$1/\gamma_I$
Contact rate (AI)	116 per thousand per day	0.00026	$c^{(AI)}$
Contact rate (VI)	112 per thousand per day	0.00028	$c^{(VI)}$
Contact rate (NS)	111 per thousand per day	0.00026	$c^{(I)}$
Avg. duration of partnership (AI)	323 days	20.6	$1/\rho^{(AI)}$
Avg. duration of partnership (VI)	556 days	46.2	$1/\rho^{(VI)}$
Avg. duration of partnership (NS)	909 days	52.8	$1/\rho^{(I)}$
Transmission probability per contact (acute infection, AI) (Leynaert et al., 1998)	18.3%	.083	$\tau_I^{(AI)}$
Transmission probability per contact (acute infection, VI) (Leynaert et al., 1998)	0.92%	0.0013	$\tau_I^{(VI)}$
Transmission probability per contact (acute infection, NS) (Kaplan and O'Keefe, 1993)	8.8%	0.088	$\tau_I^{(I)}$
Transmission probability per contact (chronic infection, AI) (Leynaert et al., 1998)	1.4%	.010	$\tau_J^{(AI)}$
Transmission probability per contact (chronic infection, VI) (Leynaert et al., 1998)	0.07%	0.0001	$\tau_J^{(VI)}$
Transmission probability per contact (chronic infection, NS) (Kaplan and O'Keefe, 1993; White et al., 2007)	0.67%	0.0067	$\tau_J^{(I)}$
Serosorting rate	0.336	–	μ

Each partnership type has its own associated contact rate, transmission probability per contact, and average duration.

Our analysis was based on the empirical distribution of contacts, which is generated by

$$g(x_A, x_I, x_V) = \sum_{ego \in S} w(ego) x_A^{d_A(ego)} x_I^{d_I(ego)} x_V^{d_V(ego)} \Big/ \sum_{ego \in S} w(ego). \quad (4)$$

where $g(\cdot)$ is a probability generating function (Wilf, 2006), and d_A, d_I, and d_V are the numbers of concurrent AI sex, needle sharing, and VI sex partners, respectively. The survey weight is denoted $w(ego)$.

Table 2 gives our model of HIV transmission through contact networks with the following properties:

- An arbitrary distribution of partnerships with extended duration
- Multiple partnership types with distinct transmission rates
- Partnerships with variable durations
- Multiple stages of infection with distinct transmission and removal rates
- Preferential attachment by serostatus (serosorting).

The modeling framework is a relatively low-dimensional system of ordinary differential equations that express the dynamics of edges between susceptibles and infecteds. Expressing the dynamics in terms of edges (as opposed to individuals or nodes) simplifies the system (reducing the number of equations needed). We here provide some intuition for how the model equations work, while Appendix A provides detailed derivations. The approach is similar to the one taken by Volz and Meyers (2007). $M_{XY}^{(i)}$ is proportional to the number of half-edges of partnership type i from a node of type X to a node of type Y. The total number of such half-edges in a network of size N is $NM_{XY}^{(i)}$. Then, for example, the probability that a given type i partnership with a susceptible goes to an acutely infected individual and is thus a potential conduit for infection is $M_{SI}^{(i)} / M_S^{(i)}$ where $M_S^{(i)}$ is proportional to the total number of half-edges connected to susceptible individuals (less by a factor of N). Then we can calculate the hazard of infection from acutely infected individuals for a susceptible node with a number of risky partnerships described by the integers d_A, d_V, and d_I. The hazard is proportional to

$$d_A \beta_I^{(A)} M_{SI}^{(A)} / M_S^{(A)} + d_V \beta_I^{(V)} M_{SI}^{(V)} / M_S^{(V)} + d_I \beta_I^{(I)} M_{SI}^{(I)} / M_S^{(I)}.$$

The survivor function for escaping infection for a single partnership of each type is then found by integrating such a hazard function; the sequence of survivor functions for each partnership type is denoted by the vector $\vec{\theta}$.

The number of susceptible and infected individuals can be retrieved from the system via the probability generating function for the number of concurrent partnerships. The probability of someone with d_A, d_V, and d_I risky partnerships remaining susceptible would be $\theta_A^{d_A}\theta_V^{d_V}\theta_I^{d_I}$. Then the fraction of the population remaining susceptible at any time is given by $S = g(\vec{\theta})$, where $g(\cdot)$ is the generating function defined in Eq. (4). Further details can be found in Appendix A.

Table 2
Model equations. S is the fraction of the population susceptible. M_{SS}, M_{SI} and M_{SJ} are proportional to the number of edges between susceptibles and other susceptibles, susceptibles and acute infecteds, and between susceptibles and chronic infecteds respectively. Each quantity has a superscript (i) which is an index for contact types (anal intercourse (AI), vaginal intercourse (VI), or sharing needles when injecting drugs (NS)). M_S, M_I and M_J are proportional to the number of half-edges such that node is susceptible, acute infected, or chronic infected respectively. T_i is a derived variable which is proportional to the number of transmissions per unit time by each contact type. $\delta(i,j)$ is a derived variable equal to the average excess degree of type i partnerships among susceptible nodes selected with probability proportional to the number of type j partnerships.

Derived variables:

$$S = g(\vec{\theta})$$
$$M^{(i)} := \left[\frac{d}{d\chi_i} g(\vec{\chi})\right]_{\vec{\chi}=\vec{1}}$$
$$M_S^{(i)} = \theta^{(i)} \left[\frac{d}{d\chi_i} g(\vec{\chi})\right]_{\vec{\chi}=\vec{\theta}}$$
$$T_i = \beta_I M_{SI}^{(i)} + \beta_J M_{SJ}^{(i)}$$
$$\delta(i,j) = \theta^{(i)} \left[\frac{d^2}{d\chi_i d\chi_j} g(\vec{\chi})\right]_{\vec{\chi}=\vec{\theta}} \Big/ \left(\left[\frac{d}{d\chi_j} g(\vec{\chi})\right]_{\vec{\chi}=\vec{\theta}}\right)$$

Ordinary differential equations:

$$\dot{\theta}^{(i)} = -\beta_I^{(i)} \theta^{(i)} \frac{M_{SI}^{(i)}}{M_S^{(i)}} - \beta_J^{(i)} \theta^{(i)} \frac{M_{SJ}^{(i)}}{M_S^{(i)}}$$
$$\dot{M}_{SI}^{(i)} = -M_{SI}^{(i)} \left(\beta_I^{(i)} + \gamma_I\right) + \frac{M_{SS}^{(i)} - M_{SI}^{(i)}}{M_S^{(i)}} \sum_j T_j \delta(i,j) - \mu^{(i)} M_{SI}^{(i)} + \rho^{(i)} \left(M_S^{(i)} M_I^{(i)} / M^{(i)} - M_{SI}^{(i)}\right)$$
$$\dot{M}_{SJ}^{(i)} = M_{SJ}^{(i)} \left(\beta_J^{(i)} + \gamma_J\right) + M_{SI}^{(i)} \gamma_I - \frac{M_{SJ}^{(i)}}{M_S^{(i)}} \sum_j T_j \delta(i,j) - \mu^{(i)} M_{SJ}^{(i)} + \rho^{(i)} \left(M_S^{(i)} M_J^{(i)} / M^{(i)} - M_{SJ}^{(i)}\right)$$
$$\dot{M}_{SS}^{(i)} = -2 \frac{M_{SS}^{(i)}}{M_S^{(i)}} \sum_j T_j \delta(i,j) + \mu^{(i)} \left(M_{SI}^{(i)} + M_{SJ}^{(i)}\right) + \rho^{(i)} \left(\left(M_S^{(i)}\right)^2 / M - M_{SS}^{(i)}\right)$$
$$\dot{M}_I^{(i)} = -\gamma_I M_I^{(i)} - \dot{M}_S^{(i)}$$
$$\dot{M}_J^{(i)} = \gamma_I M_I^{(i)} - \gamma_J M_J^{(i)}$$

Results

The prevalence of HIV in the sample is 13.3%. We estimated the joint-distribution of the number of concurrent partnerships of each type. The predominant risk factor in this population was vaginal sex (84% of partnerships), while there were very few MSM and drug-use partnerships (both comprising 8% of partnerships). Several parametric distributions (geometric, negative binomial, discretized lognormal, Poisson and power law) were fit to the data and compared using AIC. The distribution of the number of concurrent AI partnerships was geometric (mean 1.7), as was the distribution of the number of concurrent injection partnerships (mean 2.8). The distribution for the number of concurrent vaginal sex partners followed a Poisson distribution (mean 3.0). The distributions are illustrated in Fig. 3. The parametric distributions (geometric and Poisson) were not used in subsequent analysis. Rather, all models were constructed using the empirical distributions.

The epidemiological impact of different partnership types is determined by the average concurrency, transmissibility per act, the typical duration of a partnership, and the contact rate per partnership.

Contact rates per partnership were relatively constant among the three risk factors (Table 1). Partnerships between MSM had the shortest duration (323 days), while vaginal sex partnerships lasted 556 days. Needle-sharing partnerships were relatively constant, lasting 909 days on average. Seropositives and seronegatives exhibited similar numbers of concurrent partnerships, averaging 2.8 and 2.9 concurrent partners, respectively.

An important implicit assumption of the network-based models is that the contact rate and neighbor-exchange rate is specified per edge, such that people with more edges have a higher contact rate (summed over all partnerships), and the duration of a partnership is independent of the number of partners (Blower and Boe, 1993). It is possible that individuals are constrained by a finite budget of potentially risky contacts, and that heterogeneity is most pronounced in the number of potential people contacted, not in frequency of risky contact. Unfortunately, our data are not very informative about the relationship between the frequency of sex acts and degree; respondents didn't report specific acts, but rather rated the frequency of contact with each partner on a scale of 1 to 5. We did not find any evidence for a sexual budget influencing our results, though the survey may not be sensitive enough to detect one. A linear regression of mean self-reported contact rate on the number of concurrent partners did not find a statistically significant downward trend ($p = 0.43$).

We used our mathematical model to predict the dynamics of HIV through a community with these estimated contact patterns. Fig. 4

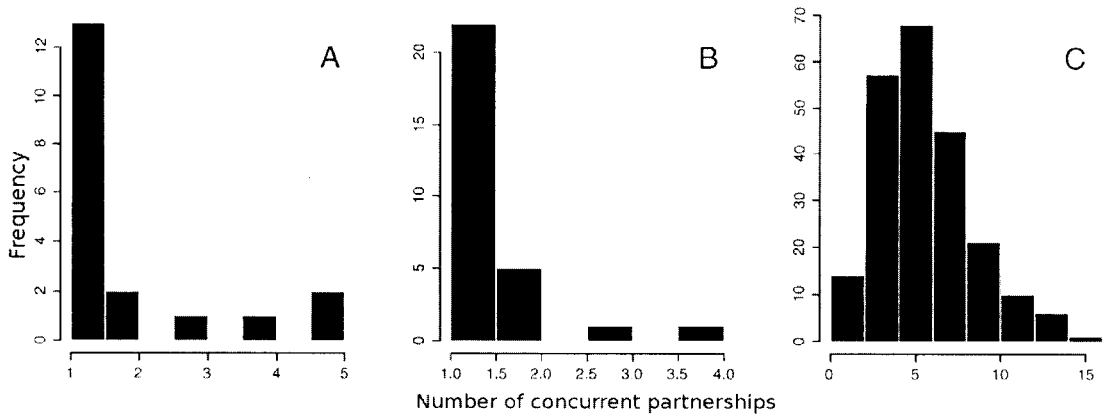

Fig. 3. Empirical distribution of the number of concurrent partners with three primary risk factors: A. anal intercourse (AI), B. needle sharing (NS), C. vaginal intercourse (VI).

shows the median and 90% confidence interval for the predicted prevalence of acute HIV infection over time. This is based on four thousand runs of the model, where each used a random sample of parameter values from the joint prior distribution of model parameters in Table 1. We model the priors as a multivariate normal distribution; standard deviations are reported in Table 1. We set the initial fraction of the population that is infected to be a normal random variable (mean = std. dev. = 10^{-3}) to reflect our poor knowledge of the actual size of the at-risk population in Atlanta. The fraction initially infected will not have a large effect on the ultimate extent of the epidemic provided the fraction is small, but it can have a large influence over the timescale of the epidemic. All initial infections started in the acute stage and were selected among those that report an AI partnership.

Substantial uncertainty about epidemic prevalence accrues from uncertainty in underlying probabilities and rates. However, the model predicts with high confidence that peak incidence will occur within about a year of the initial introduction.

To assess the impact of needle sharing, we removed needle-sharing partnerships from our model entirely. Fig. 5 compares the predicted HIV prevalence in the presence and (hypothetical) absence of transmission through needle sharing. Trajectories illustrate the median prevalence from 4000 model runs at a given time point. Parameters in each model run were drawn independently from the distributions in Table 1. The absence of needle-sharing transmission greatly diminishes the spread to the heterosexual population and decreases the number ultimately infected. 33% fewer individuals are infected after five years in the population without IDUs. But this decline is almost entirely due to infections prevented among HET. Very few infections are prevented by removing injection drug use among individuals that report at least one partnership with a PRF of AI. However IDUs have little impact on the time of peak incidence, which occurs after ≈ 225 days and depends primarily on the rapid spread among MSM.

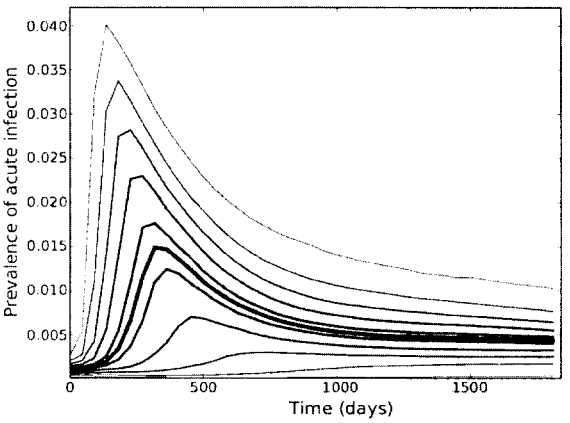

Fig. 4. Predicted prevalence of acute HIV infection over five years. The thickest line shows the median prevalence from 4000 model runs at a given time point. Each trajectory is based on a distinct set of parameters drawn independently from the set of prior distributions in Table 1. The thinnest lines at the top and bottom show 90% confidence intervals (the 5% and 95% quantiles). Intermediate lines show quantiles in increments of 10%: 45–55%, 35–65%, 25–75%, and 15–85%.

IDUs dramatically impact the fate of the epidemic, despite the fact that they are directly responsible for very few transmissions. Fig. 6 shows the fraction of transmissions that are due to infectious individuals who are acute and chronic and engaging in any of the three risk behaviors considered in the model.

The epidemiological importance of IDUs can be explained by their high centrality in the risky contact network. Betweenness centrality is a measure of how many shortest paths between pairs of individuals route through a particular individual (Freeman, 1977). The betweenness of a node *ego* is defined as the fraction of shortest paths between all pairs of actors ($alter \neq ego, alter' \neq ego$) which pass through *ego*. The average of this value tends to go down with the size of the network, so small values are unsurprising; only relative values of betweenness are important for comparing IDUs and non-IDUs. Although one must be cautious when calculating centrality from incomplete network data, we find that individuals in the Atlanta Urban network with at least one IDU contact have an average betweenness centrality of 0.0058 while non-IDUs have an average betweenness centrality of 0.0024.

The importance of IDUs is partly attributable to the high transmissibility per act of NS relative to VI (Table 1). We did a sensitivity analysis (not shown) to investigate a scenario with very low transmissibility by needle sharing. When the transmissibility by needle sharing is equal to the transmissibility by VI, we observe a reduced bridging effect by IDUs; there was lower prevalence in the heterosexual population ($\approx 13\%$ after five years), and the removal of IDUs produced a smaller decrease in the prevalence ($\approx 0.5\%$ after five years). The studies we have cited, however, indicate that the transmissibility values used in our initial analysis are realistic. Because there is so much uncertainty about the true value of this parameter, we used an uninformative prior to generate these results.

Discussion and conclusion

HIV epidemics continue to emerge in the developing world (Grassly et al., 2003; Abramovitz et al., 2009). A thorough understanding of how HIV epidemics unfold in their early stages remains important for the design of successful interventions.

Mathematical models have the potential to increase our understanding of early dynamics of HIV epidemics. For example, they can focus our attention on certain groups or behaviors that have a

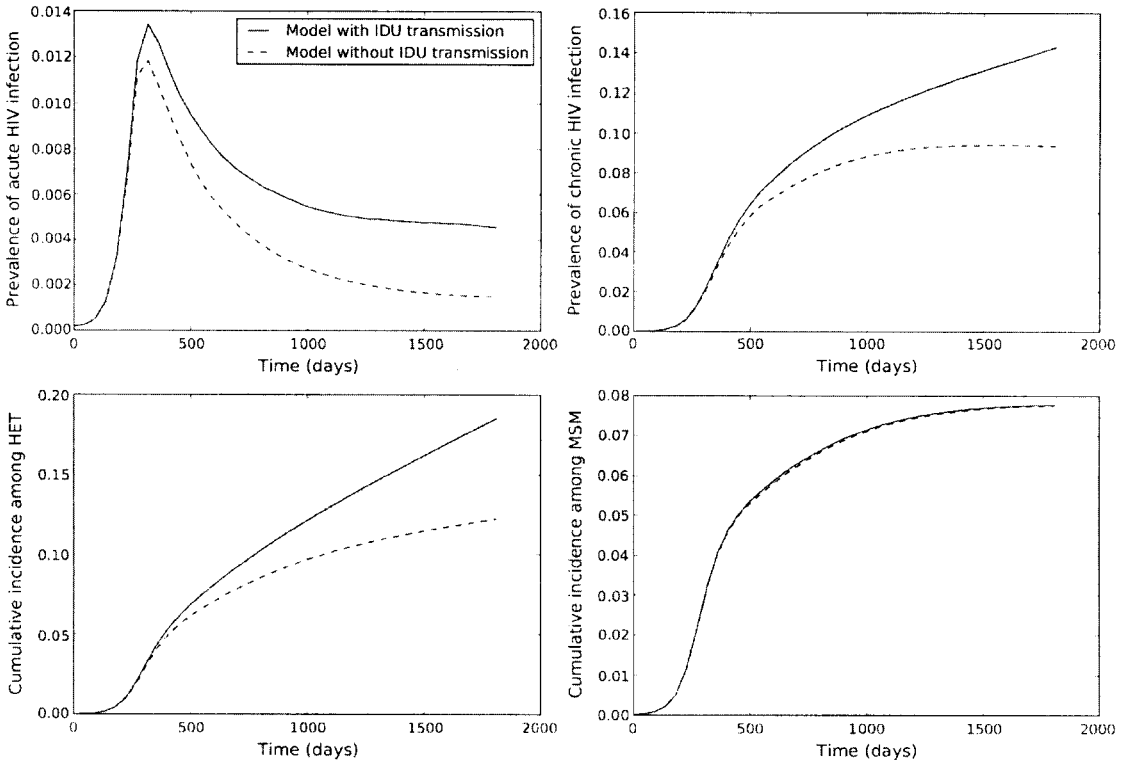

Fig. 5. Prevalence of acute and chronic infections (top) and cumulative infections (bottom) over five years. Cumulative infections are sub-categorized by primary risk factor: VI (HET) bottom left and AI (MSM) bottom right. Trajectories are shown both for the standard model (solid line) and a hypothetic model in which all transmissions via injection drug use are removed (dashed line).

disproportionate impact on the growth of the epidemic. As we have shown, the relationship between behavior, contact networks, and epidemic outcomes can sometimes be non-obvious. Although injection drug use does not account for very many transmissions in our study population, it can play an important role in determining the ultimate extent of the epidemic.

The epidemiological impact of a given risk behavior is modulated by concurrency, transmissibility per act, the duration of partnerships, and the contact rate per partnership. All of these factors can vary from individual to individual. It is difficult to incorporate such behavioral complexity into a simple mathematical model, and many different approaches have been tried. We believe that the approach we have taken strikes a good balance between realism and tractability. Incorporating as much detail as we have would often require a computationally expensive simulation, whereas we have accomplished this task within a single mathematical framework consisting of eighteen ordinary differential equations. Simplistic mass-action models would have difficulty making full use of these data, and might not reveal the complex relationship between risk behavior and epidemic outcomes that we found.

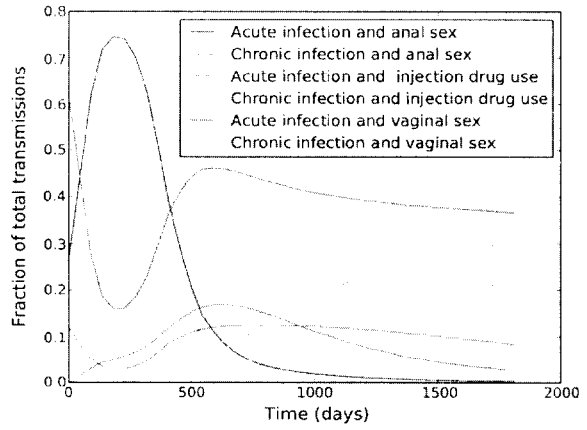

Fig. 6. The predicted fraction of transmission events that are due to anal intercourse (AI), vaginal intercourse (VI), and needle sharing (NS) over five years. These fractions are further categorized by whether transmission was committed by an acute or chronically infected individual.

Our model also motivates the collection of new kinds of data in future behavioral surveillance. Many parameters are readily estimated using chain-referral sampling data. But there are relatively few behavioral surveys that collect all of the social network information necessary for parameterizing complex network models such as the one presented here. These models require detailed information on the duration, overlap, and concurrency of partnerships as well as the diversity of these values at the individual level. Ideally, such data should also provide information on the population-scale network of contacts, so that the centrality of certain actors, such as IDUs, can be assessed.

Our data was based on longitudinal investigation which repeatedly returned to the same individuals. There was considerable attrition in later waves of the study, suggesting that diary-based sampling methods may be better for estimating the changing structure of contact networks (Mossong et al., 2008).

While our modeling approach is largely a generalization of the approach taken by Volz and Meyers (2007), a novel feature is our ability to account for preferential attachment of individuals with matching serostatus. Many network-based samples have issues related to correlation between sample units who are connected in the social network (Christakis and Fowler, 2007). This correlation may arise from the tendency for an attribute to spread from a node to a neighbor (e.g. social influence). Or it may arise from partnerships dissolving and reforming with preferential attachment (e.g. social selection). Determining the relative importance of these mechanisms

has become a notorious problem when interpreting cross-sectional network data. Our network model allows us to predict how much correlation between network partners to expect as a function of time, epidemic prevalence, and serosorting rates. Consequently, we are able estimate serosorting rates by finding the value that is most consistent with the observed partner correlations and epidemic prevalence. Although we did not find evidence for significant serosorting in our study population (Appendix B), this indirect approach is potentially generalizable to other domains including those outside epidemiology.

While our model accounts for many aspects of network structure and evolution, there are many directions in which it could be extended. For example, it considers a closed population without immigration or natural mortality, and thus it is not appropriate for modeling long time periods or populations experiencing inflows or outflows. We have not assessed the importance of insertive/receptive roles in sexual partnerships. And, we have also not accounted for certain higher-order correlations in the contact network, such as the possibility that an MSM that injects drugs may be more likely to inject with another MSM than a non-MSM. While we did not find any evidence of such correlations, this could reduce the bridging role played by IDUs to the heterosexual community. Only six out of 53 needle-sharing partnerships reported any sexual risk behaviors, and the frequency of anal intercourse in these partnerships was not significantly different from the sample frequencies (χ^2 test, $p = 10^{-7}$). These models can also be used in the future to evaluate and optimize targeted interventions such as pre- and post-exposure prophylaxis (PREP) (Lima et al., 2008; Granich et al., 2009; Lalani and Hicks, 2008). PREP has been investigated previously using mass-action compartmental models; and network models might allow a more nuanced consideration of PREP targeted at core-groups or bridging population such as IDUs.

Contributors

EMV wrote the manuscript, developed the model, designed and executed experiments, calculated survey weights, and analyzed the data. SDWF developed the model. LAM wrote the manuscript, developed the model, and analyzed the data. RR collected the data.

Acknowledgments

EMV acknowledges support from NIH T32 AI07384 and NIH R01 AI078752. SDWF was supported by the Royal Society Wolfson Research Merit Award. LAM acknowledges support from NSF DEB-0749097 and the James S. McDonnell Foundation. RR acknowledges support from NIH 1R21 DA024611-01. All authors acknowledge support of Bill and Stephanie Sick and the Santa Fe Institute.

Appendix A. Dynamic network model

Every individual has a number of contacts of different types. There will be m contact types. The probability that a random individual will have number $d_1, d_2, \ldots, d_m$ contacts of each type is described by the degree distribution $P(\vec{d})$. The number of contacts of each type of a randomly selected individual will be generated by

$$g(\vec{x}) = \sum_{\vec{d}=0,\ldots,0}^{\infty} P(\vec{d}) \prod x_i^{d_i}. \tag{5}$$

All dynamic variables have a superscript (i) which indexes the contact types (e.g. AI, VI, or NS). $\theta^{(i)}(t)$ gives the probability that an individual with one contact of type i will remain susceptible at least to time t. The fraction of the population which is susceptible is given by

$$S = g(\vec{\theta}).$$

Acutely infected individuals transmit to type i partners at the rate equal to the transmission probability per contact times the contact rate: $\beta_I^{(i)} = \tau_I^{(i)} \times c^{(i)}$. Similary $\beta_J^{(i)} = \tau_J^{(i)} \times c^{(i)}$ is the transmission rate per type-i partnership between a susceptible and a chronic infected.

If the population size is N, the total number of half-edges in the network of type i will be

$$M^{(i)} := N \times \left[\frac{d}{dx_i} g(\vec{x})\right]_{\vec{x}=\vec{1}}. \tag{6}$$

Subsequently, and without loss of generality, we will scale all variables as if $N = 1$.

$\vec{M}_{SS}$, $\vec{M}_{SI}$, and $\vec{M}_{SJ}$ will be proportional to the number of edges between susceptibles and susceptibles, susceptibles and acute infecteds, and susceptibles and chronic infecteds respectively. These variables will be normalized such that the maximum value of $M_{XY}^{(i)}$ will be $M^{(i)}$. $\vec{M}_S$, $\vec{M}_I$, and $\vec{M}_J$ will be proportional to the number of edges such that one of the two nodes is susceptible, acute infected, or chronic infected respectively.

The hazard of infection for contact type i is proportional to the number of contacts from susceptibles to infecteds, $M_{SI}^{(i)} / M_S^{(i)}$ and $M_{SJ}^{(i)} / M_S^{(i)}$, where the superscript gives the index of the contact type. This leads to the time derivative of θ (further details can be found in Volz and Meyers (2007)):

$$\dot{\theta}^{(i)} = -\beta_I^{(i)} \theta^{(i)} \frac{M_{SI}^{(i)}}{M_S^{(i)}} - \beta_J^{(i)} \theta^{(i)} \frac{M_{SJ}^{(i)}}{M_S^{(i)}}. \tag{7}$$

The equations governing the evolution of the M variables are most easily understood in terms of several derived variables.

- The number of contacts of type i made by susceptibles is

$$M_S^{(i)} := \theta^{(i)} \left[\frac{d}{dx_i} g(\vec{x})\right]_{\vec{x}=\vec{\theta}}.$$

- The transmissions per unit time by contact type i is proportional to the number of edges from susceptibles to infecteds:

$$T_i := \beta_I M_{SI}^{(i)} + \beta_J M_{SJ}^{(i)}$$

- Nodes newly infected via partnerships of type j will have a number of contacts generated by

$$\frac{d}{dx_j} g(\vec{x} \otimes \vec{\theta}) \Big/ \left[\frac{d}{dx_j} g(\vec{x} \otimes \vec{\theta})\right].$$

- The average *excess degree of type i* of a susceptible node that is newly infected via a partnership of type j is the mean of the preceding distribution, and is given by

$$\delta(i,j) := \theta^{(i)} \left[\frac{d^2}{dx_i dx_j} g(\vec{x})\right]_{\vec{x}=\vec{\theta}} \Big/ \left(\left[\frac{d}{dx_j} g(\vec{x})\right]_{\vec{x}=\vec{\theta}}\right)$$

- This is the average degree of a node infected at time t not counting the partnership which it was infected (Meyers, 2003).

To evaluate Eq. (7), we must know $M_{SI}^{(i)}$ for each contact type i. This is found by adding up the following terms:

- Contacts of type i from acute infected to susceptibles are eliminated at a rate equal to the transmission rate plus the rate of progression to the chronic phase

$$-M_{SI}^{(i)} \left(\beta_I^{(i)} + \gamma_I\right).$$

- At the rate ρ, an edge is re-wired. With probability proportional to the fraction of half-edges connected to acute infecteds, $M_I^{(i)}/M^{(i)}$ the new edge starting at a susceptible will terminate at an acute infected. This is accounted for with the addition of the mixing term

$$\rho^{(i)}\left(M_S^{(i)}M_I^{(i)}/M^{(i)}-M_{SI}^{(i)}\right).$$

- Serosorting occurs at a rate μ. At each sorting event, two partnership between a susceptible and an infected are terminated and replaced with partnerships between two susceptibles and between two infecteds. This is accounted for with the addition of the sorting term

$$-\mu^{(i)}M_{SI}^{(i)}.$$

- At time t, a number of transmissions proportional to $T_j(t)$ occur via contacts of type j.
  - Individuals newly infected via contacts of type j will have an average of $\delta(i,j)$ partnerships of type i.
  - A fraction $M_{SS}^{(i)}/M_S^{(i)}$ of the contacts from susceptibles go to other susceptibles, and will subsequently lie between susceptibles and infecteds.
  - A fraction $M_{SI}^{(i)}/M_S^{(i)}$ of these go to infecteds, and will subsequently lie between infecteds and infecteds.

  So the total change per unit time due to transmission

$$\frac{M_{SS}^{(i)}-M_{SI}^{(i)}}{M_S^{(i)}}\sum_j T_j\delta(i,j)$$

Adding these terms for all contacts of type j results in the following solution:

$$\begin{aligned}\dot{M}_{SI}^{(i)} &= -M_{SI}^{(i)}\left(\beta_I^{(i)}+\gamma_I\right)\\ &\quad+\frac{M_{SS}^{(i)}-M_{SI}^{(i)}}{M_S^{(i)}}\sum_j T_j\delta(i,j)\\ &\quad-\mu^{(i)}M_{SI}^{(i)}+\rho^{(i)}\left(M_S^{(i)}M_I^{(i)}/M^{(i)}-M_{SI}^{(i)}\right).\end{aligned}\tag{8}$$

Note that if there is only one contact type and one stage of infection, this is a much more simple equation similar to the system presented by Volz and Meyers (2007):

$$\begin{aligned}\dot{M}_{SI} &= -M_{SI}\beta+\gamma_I\\ &\quad+\frac{M_{SS}-M_{SI}}{M_S}T\delta\\ &\quad-\mu M_{SI}+\rho(M_SM_I/M-M_{SI}).\end{aligned}\tag{9}$$

The dynamics of M_{SI} are found by accounting for the following:

- Contacts of type i from chronic infecteds to susceptibles are eliminated at a rate equal to the transmission rate plus the rate of progression to the dead/inactive state, $\beta_j^{(i)}+\gamma_j$.
- As acutes move to chronic-stage infection, edges between susceptibles and acute infected change to susceptible and chronic infected at the rate $M_{SI}^{(i)}\gamma_I$.
- At the rate ρ, an edge from a susceptible to chronic infected is re-wired. With probability proportional $M_J^{(i)}/M^{(i)}$ the new edge starting at a susceptible will terminate at a chronic infected. This is accounted for with the addition of the mixing term

$$\rho^{(i)}\left(M_S^{(i)}M_J^{(i)}/M^{(i)}-M_{SJ}^{(i)}\right).$$

- Serosorting occurs at a rate γ. At each sorting event, a partnership between a susceptible and an infected is terminated and replaced with partnerships between two susceptibles and two infecteds. This is accounted for with the addition of the sorting term

$$-\mu^{(i)}M_{SJ}^{(i)}.$$

- At time t, a number of transmissions proportional to $T_j(t)$ occur via contacts of type j.
  - Individuals newly infected via contacts of type j will have an average of $\delta(i,j)$ partnerships of type i.
  - A fraction $M_{SJ}^{(i)}/M_S^{(i)}$ of these go to chronic infecteds, and will subsequently lie between acute and chronic infecteds.

  So the total change per unit time due to transmission is

$$\frac{-M_{SJ}^{(i)}}{M_S^{(i)}}\sum_j T_j\delta(i,j).$$

Adding these terms gives:

$$\begin{aligned}\dot{M}_{SJ}^{(i)} &= -M_{SJ}^{(i)}\left(\beta_J^{(i)}+\gamma_J\right)+M_{SI}^{(i)}\gamma_I\\ &\quad-\frac{M_{SJ}^{(i)}}{M_S^{(i)}}\sum_j T_j\delta(i,j)\\ &\quad-\mu^{(i)}M_{SJ}^{(i)}+\rho^{(i)}\left(M_S^{(i)}M_J^{(i)}/M^{(i)}-M_{SJ}^{(i)}\right).\end{aligned}\tag{10}$$

A similar calculation gives the dynamics of M_{SS}:

$$\begin{aligned}\dot{M}_{SS}^{(i)} &= -2\frac{M_{SS}^{(i)}}{M_S^{(i)}}\sum_j T_j\delta(i,j)\\ &\quad+\mu^{(i)}\left(M_{SI}^{(i)}+M_{SJ}^{(i)}\right)+\rho^{(i)}\left((M_S^{(i)})^2/M-M_{SS}^{(i)}\right).\end{aligned}\tag{11}$$

If there is only one contact type and one stage of infection, this is a much more simple equation similar to the system presented by Volz and Meyers (2007):

$$\begin{aligned}\dot{M}_{SS} &= -2\frac{M_{SS}}{M_S}T\delta\\ &\quad+\mu M_{SI}+\rho\left(M_S^2/M-M_{SS}\right).\end{aligned}\tag{12}$$

To account for the effects of mixing (re-wiring of partnerships), we also require the number of contacts originating from infecteds which appears in Eq. (8).

- Edges counted in $M_I^{(i)}$ change state to the chronic infected $-\gamma_IM_I^{(i)}$.
- At the rate $-\dot{M}_S^{(i)}$, edges connected to susceptibles become connected to acute infected.

$$\dot{M}_I^{(i)} = -\gamma_IM_I^{(i)}-\dot{M}_S^{(i)},\tag{13}$$

where

$$\begin{aligned}\dot{M}_S^{(i)} &= \dot{\theta}^{(i)}\left[\frac{d}{dx_i}g(\vec{x})\right]_{\vec{x}=\vec{0}}\\ &\quad+\theta^{(i)}\sum_j\dot{\theta}^{(j)}\left[\frac{d^2}{dx_ix_j}g(\vec{x})\right]_{\vec{x}=\vec{0}}.\end{aligned}\tag{14}$$

Similarly, the equation for M_j is found by noting

- Edges counted in $M_j^{(i)}$ are eliminated by recovery/death events at the rate $-\gamma_j M_j^{(i)}$.
- At the rate $\gamma_I M_I^{(i)}$, edges connected to acutes become connected to chronic infected.

Adding these terms gives

$$\dot{M}_j^{(i)} = \gamma_I M_I^{(i)} - \gamma_j M_j^{(i)}. \tag{15}$$

Appendix B. Estimation of serosorting rates

The serosorting rate, μ, is the rate at which susceptible individuals terminate partnerships with infected individuals and create new partnerships with other susceptible individuals. This is illustrated with the pseudo-chemical Eq. (3) in the main text. The behavioral survey data used in this study did not include enough information to estimate serosorting rates directly. This necessitated an indirect approach, whereby we examined a range of serosorting values, μ, within the model and selected the value which was most consistent with the observed sample HIV prevalence (13.3%) and the observed pattern of partnerships between infected and susceptible individuals.

We used the following odds ratio as a measure of network connectivity between susceptibles and infecteds:

$$OR = \frac{M_{II}/M_I}{M_{SI}/M_S}. \tag{16}$$

This is odds that a partner will be infected given that you are infected, which changes over the course of the epidemic. The observed odds ratio from the sample is 2.2.

The procedure for finding the optimal μ was the following:

1. Generate 4000 parameter sets from the multivariate normal prior distribution in Table 1. To generate sorting rates: (a) For each of the 4000 parameter sets, fix the sorting rate for NS, AI, and VI partnerships to the corresponding neighbor-exchange (or "mixing") rate times an i.i.d. uniform random variable between 0 and 5. We will call the uniform r.v. the *sorting rate expansion factor* (SREF), and results will be reported in terms of this value.

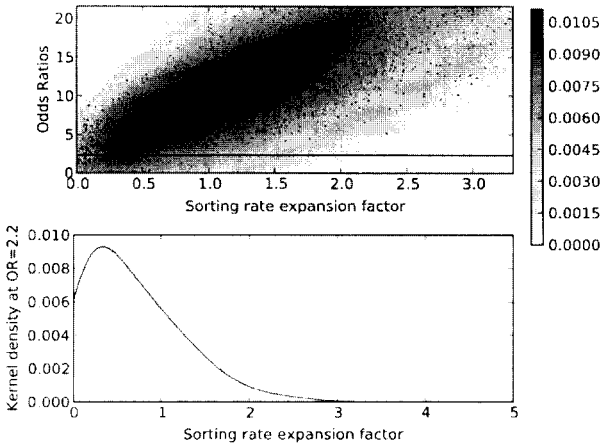

Fig. 7. Top: The odds of having an infected partner given that one is susceptible versus the ratio of the sorting rate to the neighbor-exchange rate. Each point corresponds to the solution of the model with a stochastic parameter set and at the time when the epidemic has a prevalence of 13.3%. The color scale illustrates a two-dimensional Gaussian kernel density that was fit to the cloud of points. The dashed line shows where the odds ratio equals 2.2. Bottom: The kernel density of the sorting rate when the odds ratio equals 2.2.

2. Generate 4000 model trajectories using each of the parameter sets.
3. For each trajectory, calculate the odds ratio at the point where the epidemic has a prevalence of 13.3%, which is the prevalence observed in the sample.
4. Plot the odds ratio versus the sorting rate.
5. Fit a two-dimensional Gaussian kernel density to the resulting set of points in the SREF × OR space.
6. Use the kernel density to estimate the likelihood of having a combination of 13.3% prevalence and an odds ratio of 2.2.

Fig. 7 shows the resulting collection of odds ratios and the Gaussian kernel fit. As the sorting rate increases, edges in the network become more concentrated between seroconcordant pairs (e.g. susceptible–susceptible and infected–infected), and the odds ratio increases.

The Gaussian kernel predicts that the most likely value for the SREF is quite small: 0.336. In other words, individuals choose partners on the basis of serostatus only about a third as frequently as they choose partners for other reasons.

It also predicts that there is 32.9% probability that there will be an odds ratio of 2.2 (the observed value) or less when $\gamma = 0$. Hence, we cannot reject the hypothesis that there was no serosorting in this population at all.

References

Abramovitz, D., Volz, E.M., Strathdee, S.A., Patterson, T.L., Vera, A., Frost, S.D.W., 2009. Using respondent-driven sampling in a hidden population at risk of HIV infection: who do HIV-positive recruiters recruit? Sexually Transmitted Diseases.

Anderson, R.M., May, R.M., 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press.

Baeten, J.M., Overbaugh, J., 2003. Measuring the infectiousness of persons with HIV-1: opportunities for preventing sexual HIV-1 transmission. Curr. HIV Res. 1 (1), 69–89.

Blower, S.M., Boe, C., 1993. Sex acts, sex partners, and sex budgets: implications for risk factor analysis and estimation of HIV transmission probabilities. J. Acquir. Immune Defic. Syndr. 6 1347–1347.

Butler, D.M., Smith, D.M., 2007. Serosorting can potentially increase HIV transmissions. AIDS 21 (9), 1218.

Callaway, D.S., Hopcroft, J.E., Kleinberg, J.M., Newman, M.E.J., Strogatz, S.H., 2001. Are randomly grown graphs really random? Phys. Rev. E 64 (4), 41902.

Cassels, S., Menza, T.W., Goodreau, S.M., Golden, M.R., 2009. HIV serosorting as a harm reduction strategy: evidence from Seattle, Washington. AIDS 23 000–000.

Catanzaro, M., Boguñá, M., Pastor-Satorras, R., 2005. Generation of uncorrelated random scale-free networks. Phys. Rev. E 71 (2), 27103.

Christakis, N.A., Fowler, J.H., 2007. The spread of obesity in a large social network over 32 years. N Engl J. Med. 357 (4), 370.

Freeman, L.C., 1977. A set of measures of centrality based on betweenness. Sociometry 40 (1), 35–41.

Granich, R.M., Gilks, C.F., Dye, C., De Cock, K.M., Williams, B.G., 2009. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet 373 (9657), 48–57.

Grassly, N.C., Lowndes, C.M., Rhodes, T., Judd, A., Renton, A., Garnett, G.P., 2003. Modelling emerging HIV epidemics: the role of injecting drug use and sexual transmission in the Russian Federation. China and India. Int. J. Drug Policy 14 (1), 25–43.

Hansen, M.H., Hurwitz, W.N., 1943. On the theory of sampling from finite populations. Ann. Math. Stat. 333–362.

Howard, D.L., Latkin, C.A., 2006. A bridge over troubled waters: factors associated with non-injection drug users having injection drug-using sex partners. J. Acquir. Immune Defic. Syndr. 42 (3), 325.

Kaplan, E.H., O'Keefe, E., 1993. Let the needles do the talking! Evaluating the New Haven needle exchange. Interfaces 7–26.

Keeling, M.J., Rohani, P., 2008. Modeling Infectious Diseases in Humans and Animals. Princeton University Press.

Kral, A.H., Bluthenthal, R.N., Lorvick, J., Gee, L., Bacchetti, P., Edlin, B.R., 2001. Sexual transmission of HIV-1 among injection drug users in San Francisco, USA: risk-factor analysis. The Lancet 357 (9266), 1397–1401.

Kretzschmar, M., Wiessing, L., 2008. New challenges for mathematical and statistical modeling of HIV and hepatitis C virus in injecting drug users. AIDS 22 (13), 1527–1537.

Lalani, T., Hicks, C., 2008. Does antiretroviral therapy prevent HIV transmission to sexual partners? Curr. Infectious Dis. R 10 (2), 140–145.

Lansky, A., Drake, A., Pham, H.T., 2009. HIV-associated behaviors among injecting-drug users—23 Cities, United States, May 2005–February 2006. Morb. Mortal. Wkly Rep. 58 (13), 329–332.

Leynaert, B., Downs, A.M., de Vincenzi, I., et al., 1998. Heterosexual transmission of human immunodeficiency virus: variability of infectivity throughout the course of infection. Am. J. Epidemiol. 148 (1), 88.

Lima, V.D., Johnston, K., Hogg, R.S., Levy, A.R., Harrigan, P.R., Anema, A., Montaner, J.S.G., 2008. Expanded access to highly active antiretroviral therapy: a potentially powerful strategy to curb the growth of the HIV epidemic. J. Infect. Dis. 198 (1), 59–67.

Lindenburg, C.E.A., Krol, A., Smit, C., Buster, M.C.A., Coutinho, R.A., Prins, M., 2006. Decline in HIV incidence and injecting, but not in sexual risk behaviour, seen in drug users in Amsterdam: a 19-year prospective cohort study. AIDS 20 (13), 1771.

Meyers, L.A., Newman, M.E.J., Martin, M., Schrag, S., 2003. Applying network theory to epidemics: control measures for *Mycoplasma pneumoniae* outbreaks. Emerg. Infect. Dis. 9 (2), 204–210.

Meyers, L.A., Newman, M.E.J., Pourbohloul, B., 2006. Predicting epidemics on directed contact networks. J. Theor. Biol. 240 (3), 400–418.

Molloy, M., Reed, B., 1998. The size of the giant component of a random graph with a given degree sequence. Comb. Probab. Comput. 7 (03), 295–305.

Morgan, D., Mahe, C., Mayanja, B., Okongo, J.M., Lubega, R., Whitworth, J.A.G., 2002. HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? AIDS 16 (4), 597.

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G.S., Wallinga, J., et al., 2008. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 5 (3), e74.

Newman, M.E.J., Strogatz, S.H., Watts, D.J., 2001. Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E 64 (2), 26118.

Pisani, E., et al., 2003. Sexual behavior among injection drug users in 3 Indonesian cities carries a high potential for HIV spread to noninjectors. J. Acquir. Immune Defic. Syndr. 34 (4), 403.

Rothenberg, R.B., Long, D.M., Sterk, C.E., Pach, A., Potterat, J.J., Muth, S., Baldwin, J.A., Robert III, T.T., 2000. The Atlanta Urban Networks Study: a blueprint for endemic transmission. AIDS 14 (14), 2191.

Saidel, T.J., Des Jarlais, D., Peerapatanapokin, W., Dorabjee, J., Singh, S., Brown, T., 2003. Potential impact of HIV among IDUs on heterosexual transmission in Asian settings: scenarios from the Asian Epidemic Model. Int. J. Drug Policy 14 (1), 63–74.

Salganik, M., Heckathorn, D.D., 2004. Sampling and estimation in hidden populations using respondent-driven sampling. Sociological Method. 34, 193–240.

Vanhems, P., Hirschel, B., Phillips, A.N., Cooper, D.A., Vizzard, J., Brassard, J., Perrin, L., 2000. Incubation time of acute human immunodeficiency virus (HIV) infection and duration of acute HIV infection are independent prognostic factors of progression to AIDS. J. Infect. Dis. 182 (1), 334–337.

Volz, E., Heckathorn, D.D., 2008. Probability based estimation theory for respondent driven sampling. J. Off. Stat. Stockholm 24 (1), 79.

Volz, E., Meyers, L.A., 2007. Susceptible–infected–recovered epidemics in dynamic contact networks. Proc. R. Soc. B 274 (1628), 2925.

Volz, E., Meyers, L.A., 2009. Epidemic thresholds in dynamic contact networks. J. R. Soc. Interface 6 (32), 233.

White, R.G., Ben, S.C., Kedhar, A., Orroth, K.K., Biraro, S., Baggaley, R.F., Whitworth, J., Korenromp, E.L., Ghani, A., Boily, M.C., et al., 2007. Quantifying HIV-1 transmission due to contaminated injections. Proc. Natl Acad. Sci. 104 (23), 9794.

Wiessing, L., Kretzschmar, M., 2003. Can HIV epidemics among IDUs 'trigger' a generalised epidemic? Int. J. Drug Policy 14 (1), 99–102.

Wilf, H.S., 2006. generatingfunctionology. AK Peters, Ltd.