Dynamic variation in sexual contact rates in a cohort of HIV-negative gay men

E.O. Romero-Severson[1], E. Volz[2], J. S. Koopman[3], T. Leitner[1] and E.L. Ionides[4]

November 26, 2014

[1] Los Alamos National Laboratory, Theoretical Biology and Biophysics
[2] Imperial College, Department of Infectious Disease Epidemiology
[3] University of Michigan, Department of Epidemiology
[4] University of Michigan Department of Statistics

email: eromero@lanl.gov

Abstract

HIV transmission models that include variability in sexual behavior over time have shown increased incidence, prevalence, and acute-state transmission rates for a given population risk profile. This raises the question of whether dynamic variation in individual sexual behavior is a real phenomenon that can be observed and measured. To study this dynamic variation, we developed a model incorporating heterogeneity in both between-individual and within-individual sexual contact patterns. Using novel methodology that we call iterated filtering for longitudinal data, we fitted this model by maximum likelihood to longitudinal survey data on a cohort of HIV-negative gay men. We found evidence for individual heterogeneity in sexual behavior over time. We simulated an epidemic process and found that inclusion of empirically measured levels of dynamic variation in individual-level sexual behavior bring the theoretical predictions of HIV incidence in closer alignment with reality given the measured per-act probabilities of transmission. The methods developed here provide a framework for quantifying variation in sexual behaviors that help to understand the HIV epidemic among gay men.

Introduction

Although the number of new HIV infections in the United States has remained approximately stable from 1991 to 2006, the incidence rate in young gay men has increased by 58% over the same period [1]. The increase in transmission rates is difficult to explain in the context of a virus with a low per-act probability of transmission [2, 3, 4], and the availability of effective prevention modes, such as condoms and anti-retroviral medication [5].

Vittinghoff et al. [6] found that unprotected receptive anal sex results in infection at a probability of 0.0028 per contact. Standard epidemiologic theory based on homogeneous, well-mixed populations [7] predicts that an HIV prevalence of 21% which is consistent with estimated MSM prevalence [8], could be reached only if each individual has 15 new sexual partners a year for an

average of 30 years. Such a high level of homogeneous, sustained high-risk behavior seems improbable.

The inconsistency between rising rates of man-to-man HIV transmission and observed contact rates and transmission probabilities has been addressed by decades of theoretical work including addition of risk-groups [9, 10, 11], enduring partnerships [12, 13, 14], natural history of infection [15, 16], and variability in the level of risky sexual behavior over time, termed episodic risk [10, 17, 18, 19]. All of the listed factors, excluding episodic risk, have been empirically well characterized. Resolving whether or not episodic risk is a real, measurable phenomenon will help parameterize more realistic models of HIV transmission dynamics.

A conceptually simple typology of HIV risk factors that illustrates the issues at play divides along two basic axes: population versus individual and between-units versus within-units. For example, the prevalence of intravenous drug users (an aggregate feature) in Sweden versus Latvia (between two populations) helps to explain the differential spread of HIV in those populations [20, 21]. Between-individual risk factors for the spread of HIV are well known: anal sex, unprotected sex, and birth to an infected mother. Considering all possible partners, a person who never has unprotected receptive anal intercourse is at lower risk than someone who does. However, the behaviors of individuals and whole populations are not constant over time producing within-unit variation over time.

This paper aims to illustrate a broadly applicable statistical framework for estimating both within-individual and between-individual episodic risk parameters from a longitudinal dataset of male-male sexual contacts. Our method relates an underling model of dynamic variation in sexual contact rates to aggregate count data in variable observation periods. A methodological innovation in this paper is to show how iterated filtering methods [22, 23], originally developed for investigating mechanistic models using time series data, can be applied to fit models in a longitudinal study with data consisting of many short time series. Finally, we show the influence of episodic risk on epidemic dynamics by simulating potential epidemic trajectories from a simple stochastic model.

Materials and Methods

Sexual contact data

The data in this study come from a large cohort of HIV negative gay men who reported having had either anal sex or receptive oral sex with ejaculation in the previous six months. The data were collected beginning in 1992, as described by Vittinghoff et al. [6]. The men were asked about the number and type of their sexual contacts over three follow-up interviews; 882 men completed the initial interview and three follow-up interviews for a total of 2 years of observation per person divided into 4 six-month observation periods.

We aggregated the count data into broadly relevant categories: total, anal, oral, protected, unprotected, with a partner of unknown HIV status ($\text{HIV}^{?}$), and with a partner believed to be HIV positive (HIV^{+}). For example, the anal category includes all reported instances of anal sex regardless of whether or not they were protected, insertive, or with a HIV^{+} partner. However, the oral contact reported in the study were only unprotected receptive oral sex with ejaculation. 'Total' refers to the number of any type of sexual contact.

The data do not differentiate between contacts made with long-term versus short-term partners such that we cannot disentangle the rate of partnership formation from the rate of sexual contact itself.

A descriptive model of contact rates that accounts for between-individual and within-individual variation

Romero-Severson et al. [24] identified four features of contact rate data that should be included in a statistical model of contact rates: (i) heterogeneity between individuals, (ii) heterogeneity within individuals over time, (iii) individual-level autocorrelation, (iv) secular trend over the duration of the cohort study. Figure 1 shows plots of these four factors in the data. To model these phenomena, we suppose that each individual has a latent rate $X_i(t)$ of making contacts of a specific type. Each data point, y_{ij}, is the number of reported contacts for individual i between time t_{j-1} and t_j, where $i = 1, \ldots, 882$ and $j = 1, \ldots, 4$. The unobserved process $\{X_i(t)\}$ is connected to the data through

the expected number of contacts for individual i in reporting interval j, which we write as

$$C_{ij} = \alpha^{j-1} \int_{t_{j-1}}^{t_j} X_i(t)\, dt, \tag{1}$$

where α is an additional secular trend that accounts for the observed decline in reported contacts. A basic stochastic model for homogeneous count data would model y_{ij} as a Poisson random variable with mean and variance equal to C_{ij} [25, Chapter 6]. However, the variance in the data are much higher than the mean of the data [24]. To account for this we assume that the data are negative binomially distributed [26, 27], which is a generalization of a Poisson distribution that allows for increased variance for a fixed mean, leading to the model

$$y_{ij} \sim \text{NegBin}\left(C_{ij}, D_i\right), \tag{2}$$

with mean C_{ij} and variance $C_{ij} + C_{ij}^2/D_i$. Here, D_i is called the dispersion parameter, with the Poisson model being recovered in the limit as D_i becomes large. The dispersion, D_i, can model increased variance compared to the Poisson distribution for individual contacts, but does not result in autocorrelation between measurements on an individual over time, which is observed in the data. To model this autocorrelation, we suppose that individual i has behavioral episodes within which $X_i(t)$ is constant, but the individual enters new behavioral episodes at rate R_i. At the start of each episode, $X_i(t)$ takes a new value drawn from a Gamma distribution with mean μ_X and variance σ_X,

$$X_i(t) \sim \text{Gamma}(\mu_X, \sigma_X). \tag{3}$$

To complete the model, we also assume Gamma distributions for D_i and R_i,

$$D_i \sim \text{Gamma}(\mu_D, \sigma_D), \tag{4}$$

$$R_i \sim \text{Gamma}(\mu_R, \sigma_R). \tag{5}$$

The parameters, σ_X, σ_D and σ_R control individual-level differences in behavioral parameters allowing the model to encompass a wide range of sexual contact patterns.

Figure 2 shows one possible realization of the model for a single individual. The distinction between the effects of the rate at which new behavioral episodes begin, R_i, and the dispersion parameter, D_i, is subtle since both model within-individual variability. The signal in the data about distinct behavioral episodes could be overwhelmed by a high variance in number of reported contacts resulting from a low value of D_i. Whether the data are sufficient to identify both R_i and D_i is an empirical question that we address below.

Likelihood-based parameter estimation and model selection

Our stochastic dynamic model for contacts is a partially observed Markov process (Web Material, section S1; see also 28). Likelihood-based inference was carried out using iterated filtering [22] implemented in `pomp` version 0.43-4 [29] running in R2.15.3 [30]. Iterated filtering is a Monte Carlo algorithm which computes the maximum likelihood estimate for partially observed Markov process models. Filtering is the numerical computation of estimating unobserved states and evaluating the likelihood function for a partially observed Markov process. Iterated filtering carries out multiple filtering operations using a sequential Monte Carlo filter, with perturbations in the unknown model parameters designed so that successive filtering operations converge toward the maximum likelihood estimate. Sequential Monte Carlo is a flexible nonlinear non-Gaussian filtering method, also known as the *particle filter* [31], in which the unknown distribution of the latent dynamic variables is represented by a Monte Carlo sample from this distribution (known as a *swarm* of *particles*). Successive iterations of the filtering process make successively smaller perturbations to the parameters, with the heuristic that the optimization process is cooling toward a freezing point which is theoretically guaranteed to be a local maximum of the likelihood function [23]. Increasing the number of particles and increasing the duration of the cooling process may improve the practical performance of the algorithm, while adding to the computational expense.

All optimizations were calculated by running sequential, fast-cooling particle filters (200 particles, cooling factor 0.95). At the end of each optimization the log likelihood was estimated by a running a single pass particle filter with 10^5 particles. A new particle filter was initialized at the current best parameter set and run using the same fast-cooling strategy.

Profile likelihood methods were used to compute maximum likelihood estimates and confidence intervals for each contact type with the selected model (table 1). The process of applying sequential particle filters was stopped once the profiles appeared to be unchanged by further applications of the particle filter. The maximum likelihood estimates reported in table 1 were found by fitting a loess curve to the profile likelihood and finding its maximum [32]. The confidence intervals were obtained by finding the point on the loess curve 1.92 log likelihood units to either side of the maximum likelihood estimate. The reported log likelihood was the maximum of the likelihoods for each profile for a given contact type. Nested models can be compared using likelihood ratio tests [33] and non-nested models by Aka-Ike's information criterion [34]. Either way, evidence for a parameter is strong if addition of that parameter increases the maximized likelihood by much more than one log unit.

Simulation from an simple epidemic model with time-variable contact rates

Simulations of the epidemic dynamics assuming the time-variable contact rates implied by our analysis were coded as Markov chains in R3.0.3 [30]. Individuals were assumed to be either susceptible (S) or infected (I) and select partners randomly proportional to their contact rates. At birth, individuals were assigned a contact rate $X_i \sim \text{Gamma}(\mu_X, \sigma_X)$ and an overdispersion parameter $D_i \sim \text{Gamma}(\mu_D, \sigma_D)$. To simulate the trajectory of the system we evaluated the probability of each event over a period of one month. Given that the model parameterization implies much less than one infectious contact per month, this time step is reasonable. At each step we determined in order: (1) which individuals became infected; (2) which individuals start a new behavioral episode; (3) which individuals died or were otherwise removed; (4) how many new individuals entered the system. Each susceptible individual makes $Q_i \sim \text{NegBin}\,(X_i, D_i)$ contacts per month. Of those $\text{Binomial}\left(Q_i, \frac{\eta_I \beta}{\eta_S + \eta_I}\right)$ are infective, where $\eta_S = \sum_{j \in S} Q_j$ is the total number of contacts made by susceptibles, $\eta_I = \sum_{j \in I} Q_j$ is the total number of contacts made by infecteds, β is the per-act probability of transmission, and $\text{Binomial}(n, p)$ is the binomial distribution with mean np and variance $np(1-p)$. A susceptible individual who has 1 or more infective contacts in a month becomes infected. Each individual re-draws a new $\text{Gamma}(\mu_X, \sigma_X)$ contact rate each month with probability

$1 - \exp(-\mu_R)$. Susceptible and infected individuals are removed with probability $\left(1 - \exp\{-\zeta\}\right)$ and $\left(1 - \exp\{-(\zeta + \delta)\}\right)$ per unit time respectively where ζ is the general rate of removal and δ is the death rate for infected individuals. In each unit of time a constant number of new susceptible individuals are added (i.e. the smallest integer greater than $\frac{\kappa}{\zeta}$) so that the long-run, infection-free population size is approximately κ. In all simulations, $\kappa = 3000$, $\zeta = \frac{1}{40 \times 12}$ month^{-1}, and $\delta = \frac{1}{10 \times 12}$ month^{-1}.

Results

Model selection

Table 2 shows strong evidence (>50 log units improvement in the likelihood) for each parameter excluding σ_R (between-individual heterogeneity in the length of a behavioral episode). The full model with σ_R is no better than the model without. With perfect likelihood maximization, model 6 can only have a higher maximized likelihood than model 5, since it contains model 5 when $\sigma_R = 0$. We explain the lower maximized likelihood for model 6 to be a result of the numerical difficulty of maximization for this more complex model, combined with no substantial statistical benefit from including this additional complexity. For subsequent analysis, we therefore adopted model 5 and fixed $\sigma_R = 0$.

The standard deviation of the number of contacts overestimates between-individual heterogeneity

The empirical sample mean and sample standard deviation was 1.53 month^{-1} and 3.28 month^{-1} respectively. The maximum likelihood parameters are shown in table 1 and figure 3. All contact types showed high levels of between-individual variability (high σ_X), suggesting long-term differences in individual contact rates. However, for every contact type, σ_X was lower than the empirical between-individual standard deviation. For total contacts, σ_X was 19% lower than the empirical between-individual standard deviation and the confidence interval for σ_X did not include the empirical standard deviation. This discrepancy arises because the raw standard deviation captures

both within-individual and between-individual variance. Consider a population with no between-individual variability in contact rates (in the long-run everyone has the same number of contacts), but with a high level of within-individual variance. If we observed that population for a limited period of time, we would see some individuals with apparently high contact rates and some with low contact rates. The difference between those individuals is due solely to chance variance about a shared average contact rate. Studies that use the empirical standard deviation of contact rates to model between-individual heterogeneity will overestimate its magnitude.

High levels of both between-individual and within-individual variability

The mean overdispersion (μ_D) is interpreted relative to Poisson variability. In the limit as μ_D approaches infinity, the number of observed contacts becomes Poisson distributed with rate determined by the integration of contact rates over behavioral intervals, C_{ij}. Decreasing μ_D increases the variance in the number of observed contacts without changing its mean. For each contact type (excluding positive and oral), the average degree of overdispersion is relatively small (larger values of μ_D). However, the overdispersion parameter is highly heterogeneous in the population (high σ_D) implying the co-existence of individuals with either highly stable (Poisson-like) and highly variable patterns of sexual contacts over time.

For most contact types the average time that an individual has a constant average contact rate is on the order of two years ($\mu_R = 0.04$ month^{-1}). Therefore, the statistical evidence for a rate $\mu_R > 0$ suggests that re-drawing contact rates (i.e., episodic behavior) is allowing additional within-individual variance that is not fully accounted for by μ_D and σ_D. Unlike the overdispersion described by μ_D and σ_D, this kind of behavioral change is on the appropriate timescale to correspond to structural life changes such as formation of new partnerships.

Transmission models parameterized from crosssectional data lead to different epidemic trajectories

In the counter-factual case where the dataset that we analyzed here was, in fact, a crosssectional survey (i.e. collapsed over observational periods) we could still estimate between-individual vari-

ability. In that case, we might consider the simple model where individuals are characterized by a single contact rate drawn from a population distribution. To parameterize that distribution we might consider the sample mean (1.53 month^{-1}) and sample standard deviation (3.28 month^{-1}) as estimates of μ_X and σ_X respectively. Figure 4 show a comparison of the dynamics under three alternative parameterizations to illustrate that problems might arise when considering crosssectional data alone. The curve labeled 'Homogeneous' illustrate random realizations of a stochastic model without any heterogeneity where the contact rate is set to the sample mean per month contact rate. The curve labeled 'Between Heterogeneity' illustrate random realizations of a stochastic model where the average contact rate is set to the sample mean per month contact rate and the population standard deviation of contact rates is set to the sample standard deviation. The curves labeled 'Within+Between Homogeneity' show the dynamics with the maximum likelihood parameters from the total contacts. In the context of this simple model, including episodic risk parameters estimated from the data greatly increases the prevalence of infection. After 35 years, omitting runs that died out, the average prevalence was 0.6% in the 'Homogenous' case, 15% in the 'Between Heterogeneity' case and 20% in the 'Within+Between Homogeneity' case (Fig. 4), which is closer to the 21% reported for the American gay male population [8]. Although this model is too simple to capture the full complexity of HIV transmission dynamics, it illustrates the potential role of heterogeneity within and between individuals for explaining observed HIV prevalence.

Contacts with HIV positive partners follow a unique pattern

All contact types except those made with a person known to be HIV positive had an empirical mean close to μ_X. The empirical rate of contact with a person known to be HIV positive (0.33 month^{-1} was much lower than the estimated value of $\mu_X = 1.62$ month^{-1}. This discrepancy is due to the fact that the contacts of this type showed a particularly high variability both within and between individuals. While many individuals reported consistent contact with an HIV infected person over the course of the study, many also reported changes of more than 100 contacts between 6 month observation periods. This particular combination of both stable and highly variable contact patters is difficult to model within our framework given the available data, making the results for this

contact type less reliable.

Discussion

In this paper we demonstrated application of a new general statistical framework for estimating both within-individual and between-individual variance in sexual behavior over time based on iterated particle filtering. This method provides a general approach for fitting arbitrary models to longitudinal behavioral data that could be expanded to include more complex and realistic behavioral models as more data becomes available. We showed that in even relatively short time-series we can discern the signature of significant within-individual behavioral variability. Our model is general enough to account for both within-individual and between-individual variability for all observed contact types excluding contacts made with a person believed or known to be HIV positive.

There is a large body of work studying how behavior is modified in response to changing attitudes or perceived risk [35] such as serosorting of HIV positive persons [36]. That literature is largely focused on how otherwise static behaviors change in response to various external conditions. Our paper is the first to demonstrate empirically, to our knowledge, that sexual behavior can be highly variable in general, i.e. that the assumption of static behavior is potentially problematic. Resolving the specific causes of the dynamic variation in sexual behavior we observed remains an open question.

Previous theoretical work has shown that episodic risk can greatly increase the rate of transmission especially from those recently infected [10, 19, 18]. A key effect of episodic risk observed in those papers was the maintenance of a high-risk susceptible pool that sustains the epidemic at higher levels than would be otherwise observed. In a population with only between-individual heterogeneity, the highest risk (e.g. high contact rates) individuals become rapidly infected while the remaining lower-risk individuals are more slowly infected. Episodic risk alters these dynamics by allowing all individuals to experience episodes of brief high-risk behavior leading to both elevated population risk and elevated acute-stage transmission. Episodic risk should be considered in the broader theoretical framework for understanding the rising MSM incidence in the United States.

Our analysis has important limitations that should be addressed in the collection and analysis of future datasets. First, the data are old (collection began in the early 1990s), which may limit their relevance to contemporary dynamics. However, given the wide-spread availability of new treatment options in the 1990s that changed the perception of HIV risk, having a baseline estimate of how sexual behavior changes over time will provide context for future analyses. We did not find any newer datasets that specifically reported counts of specific sex acts over time. Often, behavioral data was measured in terms of risk factors, such as questions asking if the interviewee had had unprotected anal sex at least once in some previous interval, rather than reporting specific counts. Although we suggest caution in how these results are implemented in future models, we have demonstrated that episodic risk is a real and measurable phenomenon. Future studies will help clarify the contemporary nature of episodic risk.

The method developed in this paper could be readily expanded to other populations where longitudinal behavioral data are available. The ideal dataset would record the time at each contact and with whom the contact was made with rather than aggregate counts over observation periods. Our method could be easily modified to get maximum likelihood estimates of potentially complex behavioral models with such data. Future work on episodic risk could include demographic and contextual covariates such as age, relationship status, or drug use that could be used to compile a more complete picture of how sexual behavior changes over the course of a person's life time. We recommend that future modeling studies for pathogens with long infectious periods include the potential for within-individual variability in behavioral parameters, and emphasize that, despite its cost, collecting longitudinal sexual behavior data is essential for understanding and preventing HIV.

Acknowledgments

This work was supported by National Institutes of Health (R01AI78752 to JSK and EORS, R01AI087520 to TL and EORS) and the National Science Foundation (DMS 1308919 to ELI)

References

[1] H. I. Hall, R. Song, P. Rhodes, J. Prejean, Q. An, L. M. Lee, J. Karon, R. Brookmeyer, E. H. Kaplan, and M. T. McKenna, "Estimation of HIV incidence in the United States," *JAMA*, vol. 300, no. 5, pp. 520–529, 2008.

[2] M. J. Wawer, R. H. Gray, N. K. Sewankambo, D. Serwadda, X. Li, O. Laeyendecker, N. Kiwanuka, G. Kigozi, M. Kiddugavu, and T. Lutalo, "Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda," *Journal of Infectious Diseases*, vol. 191, no. 9, pp. 1403–1409, 2005.

[3] S. D. Pinkerton, "Probability of HIV transmission during acute infection in Rakai, Uganda," *AIDS and Behavior*, vol. 12, no. 5, pp. 677–684, 2007.

[4] T. Hollingsworth, R. Anderson, and C. Fraser, "HIV-1 transmission, by stage of infection," *The Journal of Infectious Diseases*, vol. 198, no. 5, pp. 687–693, 2008.

[5] M. S. Cohen, G. M. Shaw, A. J. McMichael, and B. F. Haynes, "Acute HIV-1 infection," *New England Journal of Medicine*, vol. 364, no. 20, pp. 1943–1954, 2011.

[6] E. Vittinghoff, J. Douglas, F. Judon, D. McKiman, K. MacQueen, and S. P. Buchinder, "Per-contact risk of human immunodificiency virus tramnsmision between male sexual partners," *American Journal of Epidemiology*, vol. 150, no. 3, pp. 306–311, 1999.

[7] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, 1992.

[8] B. Leonard, *Prevalence and Awareness of HIV Infection Among Men Who Have Sex with Men–21 Cities, US 2008: A Reprint from 'Morbidity and Mortality Weekly Report'*. DIANE Publishing, 2011.

[9] B. L. Rapatski, F. Suppe, and J. A. Yorke, "HIV epidemics driven by late disease stage transmission," *Journal of Acquired Immune Deficiency Syndromes*, vol. 38, no. 3, pp. 241–253, 2005.

[10] J. S. Koopman, J. A. Jacquez, G. W. Welch, C. P. Simon, B. Foxman, S. M. Pollock, D. Barth-Jones, A. L. Adams, and K. Lange, "The role of early HIV infection in the spread of HIV through populations," *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, vol. 14, no. 3, pp. 249–258, 1997.

[11] R. M. Anderson, G. F. Medley, R. M. May, and A. M. Johnson, "A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS," *Mathematical Medicine and Biology*, vol. 3, no. 4, pp. 229–263, 1986.

[12] J. C. Miller, A. C. Slim, and E. M. Volz, "Edge-based compartmental modelling for infectious disease spread," *Journal of The Royal Society Interface*, vol. 9, no. 70, pp. 890–906, 2012.

[13] S. M. Goodreau, S. Cassels, D. Kasprzyk, D. E. Montaño, A. Greek, and M. Morris, "Concurrent partnerships, acute infection and HIV epidemic dynamics among young adults in Zimbabwe," *AIDS and Behavior*, vol. 16, no. 2, pp. 312–322, 2010.

[14] J.-H. Kim, R. L. Riolo, and J. S. Koopman, "HIV transmission by stage of infection and pattern of sexual partnerships," *Epidemiology*, vol. 21, no. 5, pp. 676–684, 2010.

[15] L. J. Abu-Raddad and I. M. Longini Jr, "No HIV stage is dominant in driving the HIV epidemic in sub-Saharan Africa," *AIDS*, vol. 22, no. 9, pp. 1055–1061, 2008.

[16] J. M. Hyman, J. Li, and E. Ann Stanley, "The differential infectivity and staged progression models for the transmission of HIV," *Mathematical Biosciences*, vol. 155, no. 2, pp. 77–109, 1999.

[17] S. J. Alam, E. Romero-Severson, J.-H. Kim, G. Emond, and J. S. Koopman, "Dynamic sex roles among men who have sex with men and transmissions from primary HIV infection," *Epidemiology*, vol. 21, no. 5, pp. 669–675, 2010.

[18] X. Zhang, L. Zhong, E. Romero-Severson, S. J. Alam, C. J. Henry, E. M. Volz, and J. S. Koopman, "Episodic HIV risk behavior can greatly amplify HIV prevalence and the fraction of transmissions from acute HIV infection," *Statistical Communications in Infectious Diseases*, vol. 4, no. 1, pp. 544–55, 2012.

[19] E. O. Romero-Severson, S. J. Alam, E. Volz, and J. Koopman, "Acute-stage transmission of HIV: effect of volatile contact rates," *Epidemiology (Cambridge, Mass.)*, vol. 24, no. 4, pp. 516–521, 2013.

[20] D. Balode, H. Skar, M. Mild, T. Kolupajeva, A. Ferdats, B. Rozentale, T. Leitner, and J. Albert, "Phylogenetic analysis of the Latvian HIV-1 epidemic," *AIDS Research and Human Retroviruses*, vol. 28, pp. 928–932, Aug. 2012.

[21] H. Skar, S. Sylvan, H.-B. Hansson, O. Gustavsson, H. Boman, J. Albert, and T. Leitner, "Multiple HIV-1 introductions into the Swedish intravenous drug user population," *Infection, Genetics and Evolution*, vol. 8, pp. 545–552, Sept. 2008.

[22] E. L. Ionides, C. Bretó, and A. A. King, "Inference for nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 103, no. 49, pp. 18438–18443, 2006.

[23] E. L. Ionides, A. Bhadra, Y. Atchadé, and A. A. King, "Iterated filtering," *Annals of Statistics*, vol. 39, pp. 1776–1802, 2011.

[24] E. O. Romero-Severson, S. J. Alam, E. M. Volz, and J. S. Koopman, "Heterogeneity in number and type of sexual contacts in a gay urban cohort," *Statistical Communications in Infectious Diseases*, vol. 4, no. 1, 2012.

[25] M. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*. Princeton, NJ: Princeton University Press, 2007.

[26] P. McCullagh and J. A. Nelder, *Generalized Linear Models*. London: Chapman and Hall, 2nd ed., 1989.

[27] C. Bretó and E. L. Ionides, "Compound markov counting processes and their applications to modeling infinitesimally over-dispersed systems," *Stochastic Processes and their Applications*, vol. 121, pp. 2571–2591, 2011.

[28] C. Bretó, D. He, E. L. Ionides, and A. A. King, "Time series analysis via mechanistic models," *Annals of Applied Statistics*, vol. 3, pp. 319–348, 2009.

[29] A. A. King, E. L. Ionides, C. Bretó, S. Ellner, and B. Kendall, "pomp: Statistical inference for partially observed Markov processes." R package, available on CRAN at www.r-project.org, 2009.

[30] R Development Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2006.

[31] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear, non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, 2002.

[32] E. L. Ionides, "Maximum smoothed likelihood estimation," *Statistica Sinica*, vol. 15, pp. 1003–1014, 2005.

[33] J. A. Rice, *Mathematical Statistics and Data Analysis*. Belmont, CA: Wadsworth, 1988.

[34] K. P. Burnham and D. R. Anderson, *Model Selection and Inference: A Practical Information-theoretic Approach*. New York: Springer-Verlag, 2nd ed., 2002.

[35] S. Funk, M. Salathe, and V. A. A. Jansen, "Modelling the influence of human behaviour on the spread of infectious diseases: a review," *Journal of The Royal Society Interface*, vol. 7, no. 50, pp. 1247–1256, 2010.

[36] J. T. Parsons, E. W. Schrimshaw, R. J. Wolitski, P. N. Halkitis, D. W. Purcell, C. C. Hoff, and C. A. GÃşmez, "Sexual harm reduction practices of HIV-seropositive gay and bisexual men: serosorting, strategic positioning, and withdrawal before ejaculation," *AIDS (London, England)*, vol. 19 Suppl 1, pp. S13–25, 2005.

Table 1: Maximum Likelihood Estimates for Nine Contact Types. Maximum likelihood estimates and confidence intervals for each contact type were based on profile likelihood computations (see Web Material). The parameters are defined in equations (1)–(5). Units for μ_X, σ_X and μ_R are month^{-1}, with μ_D, σ_D, and α being dimensionless.

.

Type	μ_X	σ_X	μ_D	σ_D	μ_R	α	Log Lik.
Total	1.75 (1.62, 2.01)	2.67 (2.35, 3.20)	3.81 (2.76, 4.59)	4.42 (3.30, 5.03)	0.04 (0.03, 0.05)	0.90 (0.87, 0.93)	-9552.1
Anal	1.49 (1.36, 1.64)	2.33 (2.05, 2.64)	3.47 (2.88, 4.01)	4.15 (3.38, 5.94)	0.05 (0.04, 0.06)	0.90 (0.87, 0.95)	-9092.2
Oral	0.23 (0.20, 0.27)	0.77 (0.65, 1.03)	1.07 (0.84, 1.32)	1.05 (0.65, 1.57)	0.02 (0.01, 0.03)	0.84 (0.76, 0.90)	-2622.7
Insertive	0.93 (0.85, 1.01)	1.64 (1.50, 1.94)	2.26 (2.00, 2.56)	2.42 (1.82, 3.35)	0.03 (0.03, 0.04)	0.90 (0.86, 0.94)	-7438.0
Receptive	0.51 (0.44, 0.56)	0.98 (0.89, 1.17)	2.14 (1.80, 2.61)	2.59 (2.12, 3.44)	0.08 (0.02, 0.04)	0.95 (0.90, 1.0)	-5688.4
Protected	1.07 (0.98, 1.18)	1.91 (1.73, 2.18)	2.64 (2.23, 3.11)	2.90 (1.85, 3.65)	0.04 (0.03, 0.05)	0.95 (0.89, 0.98)	-8179.9
Unprotected	0.40 (0.33, 0.46)	0.91 (0.77, 1.20)	1.19 (0.95, 1.59)	1.69 (1.14, 2.49)	0.04 (0.03, 0.05)	0.81 (0.76, 0.89)	-3775.2
Positive	1.62 (1.01, 2.32)	7.45 (5.21, 9.78)	0.99 (0.68, 1.98)	3.27 (1.92, 5.08)	0.09 (0.06, 0.11)	0.71 (0.64, 0.84)	-2190.6
Unknown	1.08 (0.98, 1.20)	1.67 (1.51, 1.90)	3.19 (2.74, 3.94)	3.76 (2.80, 5.86)	0.04 (0.03, 0.605)	0.94 (0.90, 0.98)	-8214.5

Table 2: Maximum Likelihood Estimates and Corresponding Likelihoods for Six Models for Total Contacts. The parameters are defined in equations (1)–(5). The models show statistically significant improvement (increasing likelihood) up though model 5. Units for μ_X, σ_X, μ_R and σ_R are month^{-1}, with μ_D, σ_D, and α being dimensionless.

Model	μ_X	σ_X	μ_D	σ_D	μ_R	σ_R	α	Log Likelihood
1	1.61	–	0.31	–	–	–	1	-10288.3
2	1.73	–	0.6	0.67	–	–	0.99	-9935.5
3	1.62	2.11	0.76	–	–	–	0.99	-9772.9
4	1.71	1.81	1.43	1.71	–	–	0.99	-9605.6
5	1.75	2.67	3.81	4.42	0.04	–	0.90	-9552.1
6	1.79	2.63	3.98	4.91	0.04	0.01	0.93	-9554.8

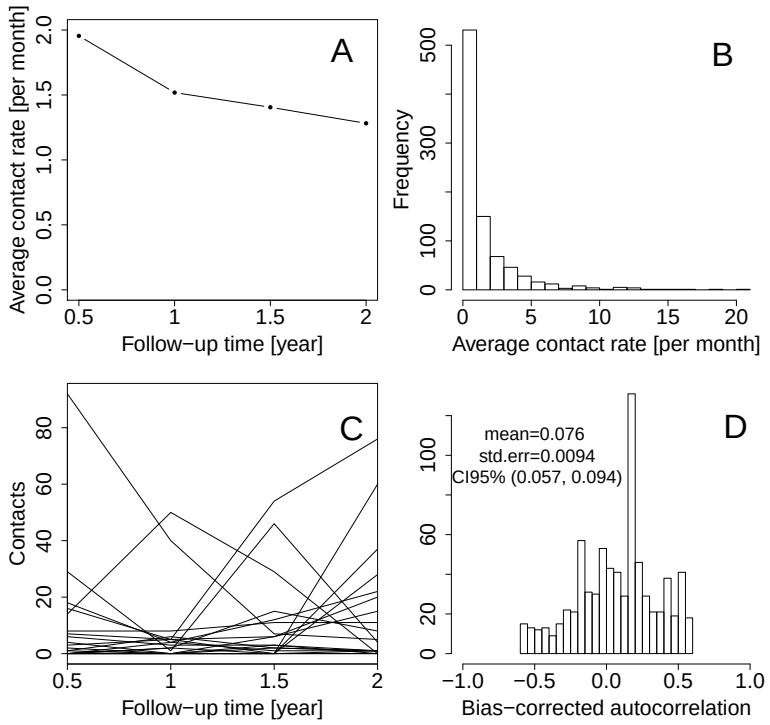

Figure 1: **Features of the longitudinal contact rate data.** (A) The secular trend sense time of enrollment into the cohort shows a clear decrease in average contact rate per person over the two-year course of the study. This is an example of a within-population change. (B) The difference in contact rates between people averaged over the course of the study demonstrates between-individual heterogeneity. (C) A time plot of reported contacts for 20 randomly selected individuals shows both within-individual and between-individual variance. (D) Sample autocorrelation, corrected for bias due to only four observation periods (Web appendix, section S3), shows a small positive correlation in contacts between observation periods.

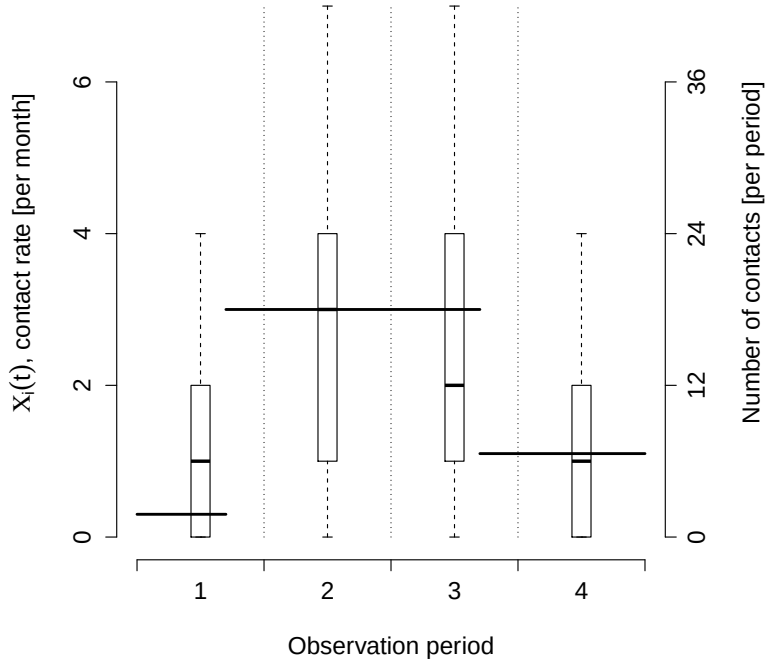

Figure 2: **Illustration of the model**. This figure presents one possible realization of the model for a single individual, i. The length of behavioral intervals (horizontal lines) over which an individual's contact rate (left axis) is constant is controlled by R_i, having mean value $1/R_i$. At the beginning of a new behavioral interval a new contact rate is drawn from a Gamma distribution mean, μ_X, and standard deviation, σ_X. The boxplots show the distribution of possible number of sexual contacts (right axis) for the given pattern of behavioral intervals shown over each 6-month long observational period. The positive latent variable D_i governs the dispersion of the actual number of contacts within a given behavioral interval. As D_i decreases toward zero, the number of contacts observed in each observation period will be more variable.

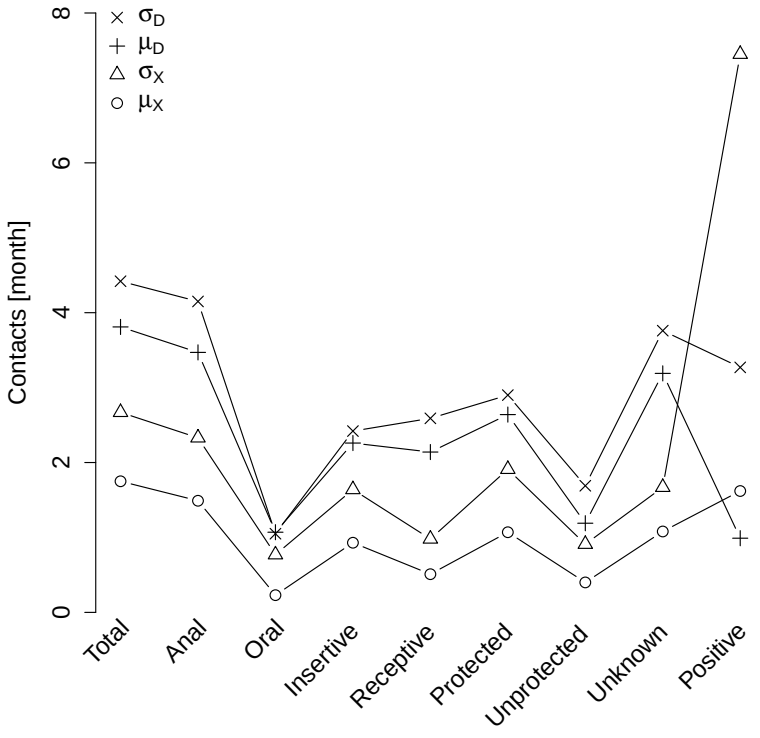

Figure 3: **Visual representation of the maximum likelihood parameter estimates.**

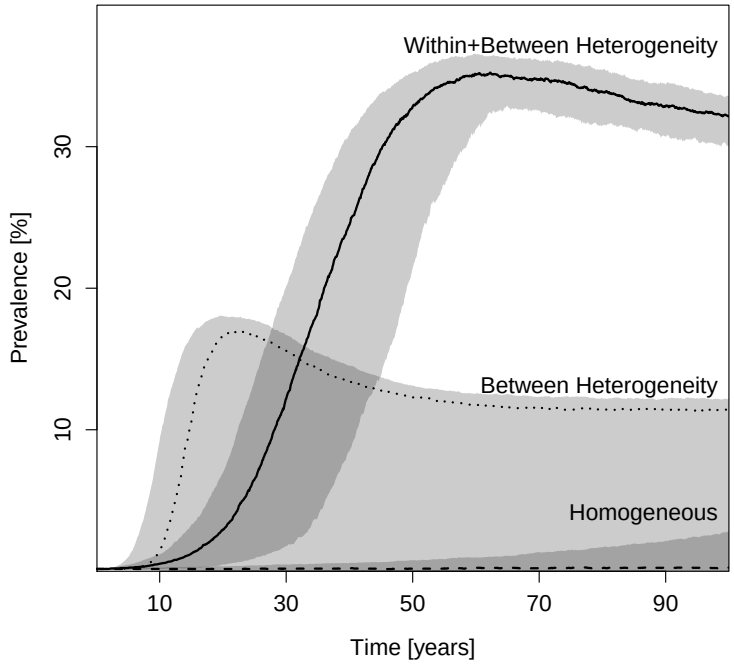

Figure 4: **Plot of epidemic trajectories under three parameterizations.** The median of 500 simulations are shown as lines and the 75^{th} and 25^{th} quantiles are shown as gray envelopes for three parameterizations. In the 'Homogeneous' case (dashed line) the epidemic was simulated where μ_X is estimated by the sample mean (1.53 month^{-1}) without any sources of between-individual or within-individual heterogeneity. In the 'Between Heterogeneity' case (dotted line) the epidemic was simulated where μ_X is estimated by the sample mean (1.53 month^{-1}) and σ_X is estimated by the sample standard deviation (3.28 month^{-1}) In the 'Within+Between Heterogeneity' case (solid line) the epidemic was simulated where each parameter is set to the estimated maximum likelihood estimate for total contacts. For both situations, the per contact probability of transmission was set to 1/120, the average length of infection was set to 10 years, and the infection-free equilibrium population size was set to 3000. The per contact probability was selected such that the basic reproduction number in the the 'Homogeneous' case was 1.53. In the 'Homogeneous', 'Between Heterogeneity', 'Within+Between Heterogeneity' cases respectively 239/500 and 172/500, 95/500 simulations died out before the 100 year mark.