The HIV Transmission Network among Men who have Sex with Men in the United States: New Insights from Dynamic Demographic Network Models

Goodreau SM, Carnegie N, Vittinghoff E, Buchbinder S

An emerging paradigm in public health seeks to tailor multiple interventions for cost-effectiveness, using epidemic modeling to identify potential areas for interaction. However, the complexities of combination interventions require new methods for multiple reasons. We report on a modeling framework developed to estimate and simulate HIV transmission among men who have sex with men. We incorporate numerous forms of demographic, relational, behavioral, and biological heterogeneity, parameterized from large-scale surveys of MSM in the United States. We rely on the ERGM framework for networks, with two novel extensions (Krivitsky and Handcock 2010, Krivitsky et al. 2011). Initial results suggest that 33% of infections occur within main partnerships, far less than the 68% estimated in a recent paper (Sullivan et al. 2009). Our estimate for the proportion of infections originating with diagnosed, untreated men is high (59%). We conclude by discussing implications of our early results, and upcoming applications to questions of combination HIV interventions for MSM.

1. Introduction

In the United States and other developed nations, HIV remains concentrated among men who have sex with men (MSM), with well over half of all new HIV diagnoses within this community (Hall et al. 2008). Rates of HIV incidence are on the rise among young MSM, and especially among young Black and Latino MSM (CDC 2009). The ecology of HIV among MSM is continually shifting, as a result of behavioral and demographic changes, changing types and patterns of HIV testing and its impact on behavior, and the gradual expansion of treatment and its impacts on survival and secondary transmission. Some aspects of this ecology – particularly those that are difficult to observe directly—remain poorly understood. These include such fundamental questions as the proportion of transmissions that occur within main partnerships vs. others, the proportion that occur from men who are aware of their positive status or not, from men who are on treatment or not, and from men in each stage of infection (acute, chronic, AIDS). Even for those features for which estimates do exist, those estimates sometimes conflict with one another, or may be out of date. One aspect of the current epidemiology that is particularly murk – and of fundamental importance – is the factors driving high incidence among young Black and Latino MSM, who consistently report individual levels of risky behavior very similar to other groups with much lower incidence.

Developing more precise answers to these questions is a key step in determining the relative potential impact of a variety of behavioral and biomedical interventions that are now or are soon to be available for MSM. This is because these interventions work on different modalities. The potential impact of couples counseling obviously depends on transmission levels within couples; prevention with positives on the degree to which transmissions occur from men who know their status; increased testing interventions on the degree to which men who don't know their status are transmitting.

The work reported in this paper represents the baseline work for a larger NIH-funded project, called PUMA (Prevention Umbrella for MSM in the Americas). Part of the MP3 (Methods for Prevention Packages Program), the goal of this work is to better assess the ways in which existing and imminent interventions can be combined, packaged and tailored for greater efficacy. Modeling is one piece of this work. The project will ultimately be modeling a wide variety of possible tailored combination interventions; here we report on the methods used to develop the model, and the baseline findings.

Because of our need to consider so many different modalities of interventions in the same model, we needed to revise existing modeling methods and philosophy considerably to make progress. General modeling philosophy, as in much of science, is to keep models as simple as is reasonable for the question at hand, in order to gain maximum clarity and interpretability. Additional detail, although increasing realism, can actually lead to less understanding; a balance point must be struck. For most modeling applications, that balance point leads to fairly simple models, but in this case, the balance point is far more complex. To be able to address the set of questions that the project is asking in any form, we require network-oriented, agent-based models. The existing tools for these models have had their own problems, however. We now lay out the existing tools, including their strengths and weakness, to motivate the methodological developments.

The most common approach traditionally for epidemic modeling involve "mass-action" or "compartmental" models expressed by systems of differential equations. However, these models require a number of simplifying assumptions that limit the kinds of questions that can be easily addressed. One particular problem is that model complexity increases multiplicatively with the number of forms of heterogeneity added into the model. Most published compartmental model have no more than five or so forms of heterogeneity among the actors in the population (e.g. age, sex, circumcision status for males, disease status, sexual activity class). The combination interventions being considered by NIH, however, require us to consider much more demographic, biological, and behavioral heterogeneity than that. At last count, the model discussed in this paper contains forty-one different demographic, behavioral, and biological attributes for each man, and that number continues to grow. Compartmental models have additional limitations – they do not allow one to consider arbitrary momentary degree distributions (that is, the proportion of men in exactly 0, 1, 2, 3 etc. ongoing relationships), which is known to affect transmission levels strongly, and is of central importance to some of our key questions. They also typically require that one discretize heterogeneity that is really continuous– for example, dividing men into those with "high", "medium" and "low" viral load.

Agent-based network models, on the other hand, offer a number of benefits – not the least of which is that their complexity only increases roughly linearly in the number of new forms of heterogeneity. Members of the population can thus be endowed with a far richer set of attributes. This framework also makes it easier to consider a wider array of hypotheses about the effects of relational structure on disease outcomes, including those pertaining to degree distributions or to forms of demographic, biological or behavioral heterogeneity that are continuous rather than discrete.

With this, and many other applications in mind, a major undertaking in the field of social network analysis for the last decade has been the development of a tractable and statistically based framework for estimating statistical models of network structure from data, conducting inference and tests of fit for those models, and simulating networks from them. The framework is known as exponential random graph models (ERGMs), and its theory and implementation have been extensively discussed in the literature (Frank and Strauss 1986, Strauss and Ikeda 1990, Wasserman and Pattison 1996, Handcock 2003, Snijders et al. 2006, Robins and Morris 2007, Hunter et al. 2008). As a result, we now have the tools for general estimation and comparison of network models of arbitrary complexity.

Use of these models has, until now, required one major tradeoff, however, which has effectively prevented their widespread application to questions of disease transmission. The versions discussed in the literature above have all required the assumption of a fixed population, and fixed values for any attributes that affect relational formation and dissolution. In other words, they have not been able to incorporate any vital dynamics or other forms of demographic change. Recent developments, discussed below, have finally overcome these limitations, extending the modeling framework into cases with dynamic sets of actors, actor attributes, and relations.

In this paper, we report on a model developed through the NIH MP3 (Methods for Prevention Packages Program) to estimate and simulate models of HIV transmission among men who have sex with men. These models incorporate demographic realism (population entry and exit, aging, race/ethnicity), relational richness (both long-term partnerships and short-term contacts), behavioral diversity (mixing patterns by race and age in both types of partnerships, rates of unprotected sex based on diagnosis and disclosure status, degree distributions), biological diversity (viral load, infectivity), treatment status, circumcision status, and many other forms of heterogeneity necessary to consider intervention packages for this population. The model is parameterized using data from multiple large-scale behavioral and

demographic surveys of MSM, including the baseline survey of the Explore study (Koblin et al. 2003, Chesney et al. 2003) and NHBS (MacKellar et al. 2007). Network model estimation uses the exponential random graph model framework found in the R *statnet* packages, modified to address relational dynamics (Krivitsky and Handcock 2010). Network and disease simulation employs an additional extension to the ERGM framework that addresses changes to model parameters as population size and composition change (Krivitsky et al. 2011).

2. Methods

We outline our overall approach first, and then describe the steps in more detail below. In reading the outline, it is important to understand that relations with positive duration are handled separately from instantaneous ("one-off") contacts, since the underlying mathematics of the two types of networks is distinct.

Step 1: Obtain a model and data-driven estimates for the dynamic main partnership network

- Step 1a. Generate estimates of cross-sectional main-partner network structure for a population of MSM using egocentric network data.
- Step 1b. Generate information on the durations of main-partner relationships for this population.
- Step 1c. Specify a "separable–temporal ERGM" (Krivitsky and Handcock 2010) that encapsulates the network processes described by 1a and 1b.
- Step 1d. Use simulated annealing to generate a complete network with the demographic composition and network structure implied by Step 1a.
- Step 1e. Estimate the model from Step 1c on the population from Step 1d, using the separable-temporal ERGM estimation methods of Krivitsky and Handcock (2010), combined with an approximation method derived by Carnegie et al. (2011).

Step 2: Obtain a model and data-driven estimates for the cross-sectional casual contact network

- Step 2a. Generate estimates of cross-sectional casual (one-off) contact main-partner network structure for a population of MSM using egocentric network data.
- Step 2b. Specify a standard cross-sectional ERGM that encapsulates the network processes described by 2a.
- Step 2c. Use simulated annealing to generate a complete network with the demographic composition and network structure implied by Step 2a.
- Step 2d. Estimate the model from Step 2b on the population from Step 2c, using standard cross-sectional ERGM methods.

Step 3: Generate demographic parameters (e.g. all-cause ASMRs, HIV mortality rates by viral load), from data. Parameters and their sources are detailed in Table 1.

Step 4. Generate additional parameters (e.g. testing, treatment, transmission by viral load) from data. Parameters and their sources are detailed in Table 1.

Step 5. Simulate each scenario over time, using the following general approach. For each time step:

- update vital dynamics
- update other attributes (testing status, viral load, etc.)

- update main network stochastically contingent on the new population structure and attributes, using the size- and composition-invariant methods of Krivitsky et al. (2011)
- model acts of unprotected anal intercourse (UAI) within main partnerships for that time step
- generate a new cross-sectional casual UAI contact network contingent on the new population structure and attributes
- model transmission events within each UAI event

We implement network simulation and estimation using statnet (http://www.statnetproject.org; Handcock et al. 2003). The remained of the tasks are coded in *R*.

Underlying both Steps 1 and 2 is the exponential random graph model (ERGM) framework discussed earlier. For Step 1, however, we must make the switch from cross-sectional ERGMs to dynamic ERGMs. Dynamic ERGMs are used to predict the probability of a partnership forming or dissolving. The dynamic ERGM model used is a separable formation and dissolution model (Krivitsky and Handcock 2010) to independently control the formation and dissolution of main partnerships. This model is essentially two parallel ERGM models, one that acts only on the set of empty dyads (formation), and another that acts only on the set of existing partnerships (dissolution). The changes in the two parallel operations are then merged to form the realization of the network at the new time step.

We specify the partnership formation part of the ERGM with an edges term to control density, an individual attribute of race to allow for differential rates of partnership formation across races, partnership-level matching on race, age and preferred sexual role, and a term for the number of men in two simultaneous relationships to control for the tendency to form concurrent partnerships. Represented in conditional auto-logistic form the model is:

$$\text{logit}P(y_{ij}=1 \mid Y_{ij}^{c}) = \Theta(e)\delta(e) + \Sigma_r\Theta(u_r)\delta(u_r) + \Sigma_r\Theta(m_r)\delta(m_r) + \Theta(a)\Sigma_{k<l}\delta(\text{abs}(\text{sqrt}(\text{age } k) - \text{sqrt}(\text{age } l))) + \Theta(d_2)\delta(d_2) + \Sigma_{c=r,v}\Theta(m_c)\delta(m_c)$$

where y_{ij} = the pair of persons *i* and *j*, and $y_{ij} = 1$ indicates they are partners; Y_{ij}^{c} = the rest of the pairs in the network, excluding the y_{ij} pair; e = total number of partnerships of all types in the network; u_r = # of partnerships of persons of race r; m_r = # of partnerships with both partners of race r; m_c = # of partnerships with both partners of role class c (which can take values i = strictly insertive; r = strictly receptive; v = versatile; see below); k and l represent the actors in each main partnership; d_2 = # of actors in exactly two main partnerships at a given time. The function δ represents the change in the specific network statistic (e, u, m or d) when the pair i,j become partners. Θ represents the coefficients for each term.

The model for dissolution of partnerships is a simple Bernoulli model with parameter determined by the average partnership duration. The parameter in this case does not control density directly (as it would in a static Bernoulli model), but rather the probability of a partnership persisting through the current time step. This, together with a Markov independence assumption implicit in the model, generates an exponential survival curve with mean approximately 2.2 years for all main partnerships.

Step 1e consists of estimating the values of the Θ coefficients that correspond to observed sexual network patterns. Estimating these parameters directly for the dynamic ERGM is extremely computationally intensive (on the order of days to weeks), and can be unstable if the starting values used for the MCMC algorithm are far from the true model parameters. We use an adjustment to the static ERGM fit for the formation model to obtain estimates of the dynamic model parameters (Carnegie

et al. 2011). Since the formation model controls only formation of new ties, using the edges term from the static ERGM fit will lead to high-density dynamic networks. We can adjust the edges term by subtracting off the dissolution parameter in order to obtain a parameter estimate that results in networks with appropriate density. This can be thought of as requiring a lower density in the sub-graph of empty dyads on which partnership formation acts in order to maintain the appropriate density in the full graph.

Note that estimation occurs on a fixed set of nodes with fixed attributes. Later, updating the network at each time step will need to occur in the context of a set of nodes that are changing, both in their set (due to births and deaths) and attribute composition (due to aging, disease transmission, testing, etc.) The approach laid out in Krivitsky et al. (2011) is what will allow us to use the estimated values we have just obtained in Step 1e for a network of fixed composition on a demographically dynamic one. We return to this point below.

Casual contact model. For the casual model, the network is defined as those pairs having acts of UAI on a given day. The specific casual model that we include is:

$$\text{logitP}(y_{ij}=1 \mid Y_{ij}^{c}) = \Theta(e)\delta(e) + \Sigma_r\Theta(u_r)\delta(u_r) + \Sigma_r\Theta(m_r)\delta(m_r) + \Theta(a)\Sigma_{k<l}\delta(\text{sqrt(age k)} - \text{sqrt (age l)}) + \Sigma_t\Theta(m_t)\delta(m_t) + \Sigma_q\Theta(q)\delta(q) + \Sigma_{c=r,v}\Theta(m_c)\delta(m_c)$$

where y_{ij} = the pair of persons *i* and *j*, and y_{ij} = 1 indicates they are partners; Y_{ij}^{c} = the rest of the pairs in the network, excluding the y_{ij} pair; e = total number of partnerships of all types in the network; u_r = # of partnerships of persons of race r; m_r = # of partnerships with both partners of race r; m_c = # of partnerships with both partners of role class c (which can take values i = strictly insertive; r = strictly receptive; v = versatile; see below); m_t = # of partnerships with both partners of diagnosis status t; q = # of partnerships of persons of casual activity class q. The function δ represents the change in the specific network statistic (e, u, m or d) when the pair i,j become partners. Θ represents the coefficients for each term. Unlike in the main partnership model, here there is no constraint on degree distribution. An individual can have casual UAI with more than one person per day, with the degree distribution flowing from the rest of the parameters, and thus approximately Poisson. We explain the casual activity classes below.

With estimation for both the main and casual models complete, we can turn to Step 5, simulation. We now detail specific features of the simulation.

Vital dynamics. Entries into the population occur at a fixed rate. All men enter seronegative and at age 18. Deaths/departures occur in three ways: (1) Through non-AIDS-related deaths: these occur daily at rates derived from CDC life tables (see Table 1); (2) Through AIDS mortality. For more information, see the sections *Viral load* and *Treatment* below; (3) Through departure from our population of interest, which occurs at completion of age 65.

Casual activity class. It is not accurate to assume that men are all homogeneous with regard to their propensity to engage in casual UAI; the variance in the distribution of number of acts over the last year in all of our data sources makes this clear. That said, we have little information on the continuity in men's UAI risk from year to year over their lives, and the degree to which men transition through levels of risk. In order to maintain the high variance in the casual UAI behavior distribution, we partition men into five activity classes, each of which has a mean underlying daily risk of casual UAI equal to the mean in the corresponding quintile of the observed data.

Role. Men are assigned one of three roles, whose prevalence is derived from data (Table 1): exclusively insertive, exclusively receptive, and versatile. Men who are versatile have a preferred preference for insertivity that varies over the range (0,1) and which is drawn for each man with a uniform distribution. Both our main and casual models assign a coefficient of $-\infty$ to the two m_c terms, effectively preventing pairs of men who are both exclusively insertive or both exclusively receptive from pairing. When two versatile men pair, the two men are chosen to be the insertive partner for each contact separately, with probability r1/(r1+r2) and r2/(r1+r2), where r1 and r2 are their insertivity preferences, respectively.

Viral load. In the absence of treatment we model daily viral load as a six-parameter curve, following parameters derived from Little (1999). Viral load here is measured in log10 units:

Days 0-21:	rises linearly from 0 to 6.886
Days 21-40:	declines linearly from 6.886 to 4.5
Days 40-3650:	equals 4.5
Days 3650- 4380:	increases linearly from 4.5 to 7.0
Day 4380:	death

This corresponds to an onset of AIDS after 10 years, and death after 12 years (Buchbinder et al. 1994). Viral load follows different trajectories for those men on treatment; see treatment below.

*Transmission***.** Transmission is a function of viral load, sexual role (insertive, receptive) of the negative partner, and circumcision status if the negative partner is insertive. Estimates for transmission by viral load for each of the three possibilities (negative =insertive circumcised, negative = insertive uncircumcised; negative = receptive) were derived by using Wilson et al. (2008) provide the functional form linking viral load and transmission, but setting base transmission for a chronically infected person (assumed to have viral load 4.5) for URA and UIA from Vittinghoff et al. (1999). Since Vittinghoff did not split out their estimates by circumcision status, we assume that the circumcised:uncircimcized ratio of of 0.4 for vaginal intercourse (Gray et al. 2007) holds for anal intercourse as well, and assumed that Vittinghoff's population reflected the overall circumcision prevalence of the adult male US population. Collectively this yields estimates for probability of transmission of:

0.126% for a circumcised man during UIAI with a male positive with viral load of 4.5
0.314% for an uncircumcised man during UIAI with a male positive with viral load of 4.5
0.820% for any man during URAI with a male positive with viral load of 4.5

each of which can be generalized to contact with a man of arbitrary viral load using the functional form from Wilson (2008).

*UAI within main partnerships***.** Main partnerships may see UAI on some days and not others, which we model as a function of diagnosis status of the members, disclosure status and AIDS status. When main couples disclose, they are assumed to always disclose their status as they know it (i.e. those who have ever tested positive report so, all others report as having never tested positive.) We assume that 90% of main couples disclose (see Table 1). There are four different UAI probabilities for men: undisclosed, disclosed as serodiscordant, disclosed as seroconcordant negative, disclosed as seroconcordant positive. The last is mathematically irrelevant, since in all such cases both men are truly positive, and cannot transmit. Daily probability of UAI for the other three are derived in Table 1.

Treatment. Once men are diagnosed and have been infected long enough to have a CD4 count that merits treatment, they may begin treatment. From here, we model three treatment trajectories: on treatment and fully suppressed, on treatment and partially suppressed, never receiving treatment. Although in reality individual men may move among these three states, we have essentially no data at all on the frequency with which they do. Instead, we model individual men entering and staying on one path, in such proportions as to capture the prevalence of the three states in the population from data.

We want to know the incidence of each (that is, what proportion of men will go into each category, by race). This is not equal to the prevalence of each group among those who are HIV+, and at or beyond the time for going on treatment, because each group has a different mean life expectancy. Our data should come on two forms: (1) the proportion of MSM dying of AIDS who have never received any treatment, and (2) of those men on continuous treatment for six months or more, what proportion are fully suppressed, in the cross-section. The former is akin to an incidence measure, and the latter a prevalence measure, so they must be treated separately. Specifically, the former can be taken as an estimate of the proportion of men who will go into the never-treated trajectory, while the latter cannot be taken as an estimate for the proportion of treaters going into the partial treatment trajectory. We use estimates for the expected duration of each stage to back-calculate the incidence of each in order to match observed prevalence (Table 1).

For those men going on treatment, it does not automatically being at the time when CD4 counts would be expected to reach levels recommended for treatment by current guidelines. Instead, it varies by race based on data on CD4 count at actual treatment initiation from Swindells et al. (2002), which can be converted into expected times from infection to treatment using the functional form of Lyles et al. (2000). This yields expected time from infection to treatment initiation of:

Blacks = 4.1 years
Latinos = 5.0 years
Other = 3.6 years

Network evolution. As written, the dynamic ERGM assumes a fixed set of nodes in the network. In order to realistically model demographic processes acting on the population as the network evolves, we include an offset term to adjust for network size in the formation model (Krivitsky et al. 2011). This simple adjustment to the model parameters yields a model that maintains mean degree as network size increases, rather than density, as would occur if we used the parameters generated on the original network without adjustment. This is clearly a better approach for modeling sexual partnerships; as a general rule, we do not see that men have more partners on average in larger populations, as would be the case with constant density, but rather that the number of partners stays roughly the same regardless of network size, as is true with a constant mean degree. This approach not only maintains mean degree, but also maintains the expected mixing structures and degree distribution.

3. Results

Our model generates epidemics that map qualitatively onto the observed overall prevalence rates in this population. HIV prevalence in the latest round of NHBS was 25%; in our model, after the initial period of burn-in, prevalence equilibrates at around 27%. As we then simulate forward in time from the present, individual runs predict prevalence of 27-30% prevalence in the coming decades (Figure 2) if there were no behavioral and biomedical changes. The one exception to this is the prevalence of circumcision among men coming of age, which is lower than it is among men already in adulthood, since neonatal

circumcision rates have been falling over the past decades (see Table 1). This pattern is included in the model, and is responsible for the very slight rise in prevalence predicted over the coming decades.

Main vs. casual partner. Our models consistently estimated that main partner UAI is the source of 37% of infections (Figure 3a), despite constituting about 50% of the acts of UAI in the population. This number is considerably lower than the recent estimate of 68% by Sullivan et al. (2009). Since epidemic dynamics can be highly non-linear, we ran counterfactual models to consider the impact of reducing levels of UAI within main partnerships. Eliminating all main partner UAI reduces incidence by 43%, with the effects on incidence of various intermediate levels of reductions between 0% and 100% being in fact strongly linear (Figure 4).

Stage, Diagnosis, role and treatment (Figure 3b-e). The proportion of infections stemming from men with diagnosed but untreated infections was high. In our model, this represented 59% of new infections, more than those from undiagnosed (14%) and treated (27%) combined. This is in concert with recent models that suggests that universal treatment has great potential to decrease HIV incidence for MSM (Blower et al. 2000, Law et al. 2001, Velasco-Hernandez et al. 2002, Lima et al. 2008). These earlier works did not provide estimates for the proportion of infections by stage or by diagnosis status or treatment status, however, so the results cannot be directly compared. Nevertheless, our initial results reflect a qualitatively similar finding from these various less richly parameterized models. On the flip side, it directly conflicts with phylogenetic studies that examine clusters of outbreaks, which suggests that 30-50% of infections may occur in the acute phase, when virtually all men are undiagnosed (Yerly et al. 2001, Pao et al. 2005, Brenner et al. 2007).

Race. Our model only found modest differences in HIV prevalence and incidence by race/ethnicity, despite including racial differences in sexual mixing, circumcision rates, treatment initiation, and treatment response. Prevalence by race/ethnicity is plotted in Figure 5. Average annual incidence was Black 2.43%, Latino 2.16%, other 2.11%. These difference are nowhere near as high as the massive differences in incidence observed by the CDC (2009).

4. Discussion

The proportion of infections occurring within main partnerships is likely to be substantially lower than that found in recent, high-profile work by Sullivan et al (2009). That paper estimated that 68% of new HIV infections occur within main partnerships, a number that has upended some of the conventional wisdom of HIV prevention practice. However, the Sullivan et al. model used a stochastic, Bernoulli model, whose structure appears to imply mathematically that a single positive individual could infect their partner more than once. Our model, in contrast, is a fully specified demographic dynamic model, and estimates that only 33% of infections occur within main partnerships. We do not find any evidence for non-linear, synergistic effects of potential main partner interventions.

Our initial baseline results suggest that approaches to earlier treatment, if they can effectively deal with issues of toxicity and resistance, have the potential to dramatically reduce secondary transmission, potentially to the point of near-elimination of the epidemic among MSM. Far more modeling work must be done to confirm this initial impression, however. Most importantly, the divergent impressions created by the behavioral- and modeling-based research and the phylogenetics-based research must be reconciled.

The model demonstrates that observed racial/ethnic differences in sexual mixing, circumcision rates, treatment initiation, and treatment response explain only a very little portion of the HIV disparities among Blacks, Latinos and other populations in the US. Collectively, these phenomena have been discussed as possible sources for the disparities, but no work before now (to our knowledge) has combined them to demonstrate their joint effects. Far more work must be done to explore the topic of the sources of racial and ethnic disparities among MSM in the US. Leading topics that we hope to explore next include patterns of concurrency by race and differential age mixing within and across race/ethnicity groups. With all of these included, we hope to model the differences in mean levels of risky sex needed to generate the observed disparities, to see if these can in any way be reconciled with existing reports.

We stress that these are preliminary results. Much work has gone into the development of the model and the analysis of data in order to parameterize that model. These results are thus only the very beginning of a far more extensive analysis of the model expected to last for 2-3 years as a minimum. Most pressing is that for these baseline measures, sensitivity analyses will be conducted over those model inputs about which we have the least confidence. Once this is done, the model will be used to explore individual behaviors and interventions (circumcision, couples testing, NAAT testing, pre-exposure prophylaxis, serosorting with positives), as well as combinations of these interventions.

Although the results are preliminary, the work so far has clearly demonstrated one methodological point: epidemic modelers no longer need to trade off looking at realistic network structures with looking at realistic demographics. We now have the ability to integrate the two in a general, inferential framework for understanding the complex ecologies of disease transmission in realistic, demographically and behaviorally diverse communities.

References

Blower SM, Gershengorn HB, Grant RM. A tale of two futures: HIV and antiretroviral therapy in San Francisco. Science 2000; 287:650–654.

Brenner BG, Roger M, Routy JP, Moisi D, Ntemgwa M, Matte C, et al., editors. High rates of forward transmission events after acute/early HIV-1 infection. Univ Chicago Press; 2007.

Carnegie NB, Hunter DR, Goodreau SM. Cross-Sectional Approximations to Separable Temporal ERGM Parameters. 2001. Sunbelt XIII: International Network for Social Network Analysis, St. Pete Beach, FL.

Centers for Disease Control and Prevention. HIV Incidence, 2006. Published 2009. http://www.cdc.gov/hiv/topics/surveillance/resources/slides/incidence/slides/HIV-Incidence.pdf

Chesney MA, Koblin BA, Barresi PJ, Husnik MJ, Celum CL, Colfax G, Mayer K, McKirnan D, Judson FN, Huang Y, Coates TJ, and the EXPLORE Study Team. An Individually Tailored Intervention for HIV Prevention: Baseline Data From the EXPLORE Study Am J Public Health 2003 93: 933-938.

Chu H, Gange, SJ, Li X, Hoover DR, Liu C, Chmiel JS, and Jacobson LP. The Effect of HAART on HIV RNA Trajectory Among Treatment-naïve Men and Women: A Segmental Bernoulli/Lognormal Random Effects Model With Left Censoring. Epidemiology 2010; 21(4) S25-S34.

Frank O, Strauss D. Markov graphs. J Am Stat Assoc. 1986; 81(395):832–42.

Gray RH, Kigozi G, Serwadda D, Makumbi F, Watya S, Nalugoda F, Kiwanuka N, Moulton LH, Chaudhary MA, Chen MZ, Sewankambo NK, Wabwire-Mangen F, Bacon MC, Williams CF, Opendi P, Reynolds SJ, Laeyendecker O, Quinn TC, Wawer MJ. Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial. Lancet 2007; 369(9562):657-66.

Hall HI, Song R, Rhodes P, Prejean J, An Q, Lee LM, Karon J, Brookmeyer R, Kaplan EH, McKenna MT, Janssen RS for the HIV Incidence Surveillance Group. Estimation of HIV Incidence in the United States. JAMA, August 6, 2008;300(5):520

Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M. Statnet: software tools for the representation, visualization, analysis and simulation of social network data. J Stat Softw. 2008;24(1):1–11.

Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M. Statnet: software tools for the statistical modeling of network data. 2.0 ed2003.

Handcock MS. Statistical models for social networks: degeneracy and inference. In: Breiger RL, Carley KM, Pattison P, editors. Dynamic social network modeling and analysis: workshop summary and papers. Washington: National Academies Press; 2003. p. 229–40.

Hunter DR, Goodreau SM, Handcock MS. Goodness of fit of social network models. J Am Stat Assoc. 2008;103(481):248–58.

Koblin BA, Chesney MA, Husnik MJ, Bozeman S, Celum CL, Buchbinder S, Mayer K, McKirnan D, Judson FN, Huang Y, Coates TJ, and the EXPLORE Study Team. High-Risk Behaviors Among Men Who Have Sex With Men in 6 US Cities: Baseline Data From the EXPLORE Study. Am J Public Health 2003 93: 926-932.

Krivitsky PN, Handcock MS, Morris M. Adjusting for network size and composition effects in exponential-family random graph models, Statistical Methodology (2011), doi:10.1016/j.stamet.2011.01.005

Krivitsky PN, Handcock MS. A Separable Model for Dynamic Networks. arXiv:1011.1937v1 [stat.ME]. Submitted on 8 Nov 2010.

Krivitsky PN. Statistical models for social network data and processes. Seattle: University of Washington; 2009.

Law MG, Prestage G, Grulich A, et al. Modelling the effect of combination antiretroviral treatments on HIV incidence. AIDS 2001; 15:1287–1294.

Lima VD, Johnston K, Hogg RS, et al. Expanded access to highly active antiretroviral therapy: a potentially powerful strategy to curb the growth of the HIV epidemic. J Infect Dis 2008; 198:59–67.Little SJ,

McLean AR, Spina CA, Richman DD, Havlir DV. Viral Dynamics of Acute HIV-1 Infection. The Journal of Experimental Medicine 1999; 190(6): 841–850.

Lyles RH, Munoz A, Yamashita TE, Bazmi H, Detels R, Rinaldo CR, Margolick JB, Phair JP, and Mellors JW. Natural History of Human Immunodeficiency Virus Type 1 Viremia after Seroconversion and Proximal to AIDS in a Large Cohort of Homosexual Men. J Infect Dis 2000; 2000;181:872–80.

MacKellar DA, Gallagher K, Finlayson T, Sanchez T, Lansky A, Sullivan P. Surveillance of HIV risk and prevention behaviors of men who have sex with men---a national application of venue-based, time-space sampling. Public Health Rep 2007;122:39--47.

Pao D, Fisher M, Hue S, Dean G, Murphy G, Cane PA, et al. Transmission of HIV-1 during primary infection: relationship to sexual risk and sexually transmitted infections. AIDS. 2005; 19(1):85–90.

Robins G, Morris M. Advances in exponential random graph (p*) models. Soc Networks. 2007;29(2):169–72.

Snijders TAB, Pattison PE, Robins GL, Handcock MS. New specifications for exponential random graph models. Sociol Methodol. 2006;36(1):99–153.

Strauss D, Ikeda M. Pseudolikelihood estimation for social networks. J Am Stat Soc. 1990;85(409):204–12.

Sullivan PS, Salazar L, Buchbinder S, and Sanchez TH. Estimating the proportion of HIV transmissions from main sex partners among men who have sex with men in five US cities . AIDS 2009; 23:1153-1162.

Swindells S, Cobos, DG, Lee N, Lien EA, Fitzgerald AP, Pauls JS, and Anderson JR. Racial/ethnic differences in CD4 T cell count and viral load at presentation for medical care and in follow-up after HIV-1 infection. AIDS 2002; 16(13) 1832-1834.

Velasco-Hernandez JX, Gershengorn HB, Blower SM. Could widespread use of combination antiretroviral therapy eradicate HIV epidemics? Lancet Infect Dis 2002; 2:487–493.

Vittinghoff E, Douglas J, Judson F, McKirnan D, MacQueen K, Buchbinder SP. Per-Contact Risk of Human Immunodeficiency Virus Transmission between Male Sexual Partners. Am J Epidemiol 1999; 150(3):306-311.

Wasserman S, Pattison P. Logit models and logistic regressions for social networks. I. An introduction to Markov graphs and p*. Psychometrika. 1996;60:401–25.

Wawer MJ, Gray RH, Sewankambo NK, Serwadda D, Li X, Laeyendecker O, Kiwanuka N, Kigozi G, Kiddugavu M, Lutalo T, Nalugoda F, Wabwire-Mangen F, Meehan MP, Quinn TC. Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. Journal of Infectious Diseases 2005; 1403–1409.

Wilson DP, Law MG, Cooper D et al. Relation between HIV viral load and infectiousness: a model-based analysis. Lancet 2008; 372:320.

Yerly S, Vora S, Rizzardi P, Chave JP, Vernazza PL, Flepp M, et al. Acute HIV infection: impact on the spread of HIV and transmission of drug resistance. AIDS. 2001;15(17):2287–92.

Table 1: Data sources

The two main data sources for behavioral parameters are the baseline data from the Explore study (Koblin et al. 2003, Chesney et al. 2003), and Round 1 of the National HIV Behavioral Surveillance system (MacKellar et al. 2007). Each has been described in detail in numerous additional studies and analyses.

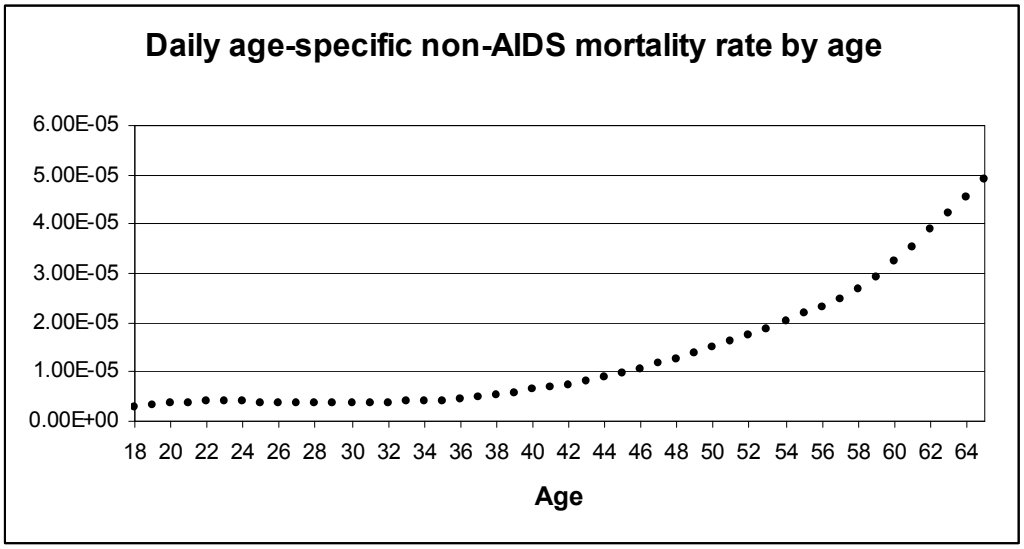
Model parameter(s)	Source(s) and value(s)
Age-specific Non-AIDS mortality rates	All –cause mortality: National Vital Statistics Reports Volume 56, Number 9, December 28, 2007. United States Life Tables, 2004 by Elizabeth Arias, Ph.D., Division of Vital Statistics, pub. (http://www.cdc.gov/nchs/data/nvsr/nvsr56/nvsr56_09.pdf) AIDS mortality (to subtract from all-cause in order to get non-AIDS mortality): National Vital Statistics Reports Volume 58, Number 8, December 23, 2009. Deaths: Leading Causes for 2005 by Melonie Heron, Ph.D., and Betzaida Tejada-Vera, B.S., Division of Vital Statistics (http://www.cdc.gov/nchs/data/nvsr/nvsr58/nvsr58_08.pdf)
Daily probability of casual UAI	Explore Placed in five quintiles: 0.0020 0.0129 0.0292 0.0610 0.2246
Race mixing in casual	NHBS. Relative log-odds: -11.53 Black-Black -12.38 Latino-Latino -12.41 Other-Other -12.72 Black-Latino -12.73 Latino-Other -13.49 Black-Other
Race mixing in steady	NHBS. Relative log-odds: - 9.63 Black-Black -10.44 Latino-Latino -10.58 Other-Other -10.89 Latino-Other -11.11 Black-Latino -11.55 Black-Other
Daily probability of UAI within an existing main partnership, by disclosure status	Explore. undisclosed 0.114 disclosed + - 0.108 disclosed - - 0.167

Diagnosis status mixing in casual contacts	NHBS. Relative log-odds: -11.67 Pos-pos -12.56 Neg-neg -13.02 Neg-pos
Age mixing within main partnerships	NHBS. Mean difference in the absolute value of the square root of the ages of main partners = 0.61.
Age mixing within casual contacts	NHBS. Mean difference in the absolute value of the square root of the ages of casual contacts = 0.73.
Age distribution	Initially set as uniform from 18-65; burnin period ensures that the age distribution settles down to what is appropriate for the given birth and death rates as well as the HIV transmission and mortality experience. Model exploration begins after the end of 100 years of burnin.
Circumcision status for population coming of age prior to the present	The STEP study (Buchbinder, personal communication) Black=85% Latino=50% Other=91% HPTN gave very similar numbers: O=88%; B=90%; L=47%
Circumcision status for population coming of age after the present	National Hospital Discharge Survey Black=70% Latino=50% Other=60%
Prevalence of role exclusivity	Explore. Reported role over previous 18 months. 8.0% exclusively insertive 8.0% exclusively receptive 84.0% versatile
Proportion of population in 0, 1, or >1 main partnerships in the cross-section	*Project T, Explore, NHBS-SF, HPTN039* None of the main studies asked this question directly. Project T (Liu, personal communication) found that 40% of men in a main partnership on the day of the interview. Neither Explore nor NHBS-SF nor HPTN039 asked the questions directly. However they all allowed one to put an upper and lower bound on the number. Upper = % with any main partner in last x months. lower = living with partner. All three center around the 40% figure as well: Explore: 24.4% to 50.2% NHBS-SF: 34.1% to 43.0%

	HPTN039: 26.4% to 54.2% Percent in >1: nobody asked this directly. However, it is important to not simply forbid it (which would underestimate the importance of main partners as men themselves define them), nor to leave it unparameterized (which would result in something like 16% of men having >1 main partner, clearly too high). Explore asked questions that allowed us to estimate upper and lower bounds as 2.8% and 0.6% of the population in 2 main partnerships at any given time. Lacking any other information, we simply took the average (1.7%).
Race composition	NHBS 17.5% Black 27.0% Latino 55.5% Other
Main partnership mean duration	1120 days NHBS
Daily probability of HIV test	1/351 Golden reports a median interval of 243 days. Assuming a geometric distribution, this corresponds to a mean interval of 351 days and a daily average probability of testing of 1/351 days. This should correspond to men getting tested almost exactly once a year on average, which is what other studies have reported qualitatively.
HIV test window period	22 days
Reduction in casual UAI once AIDS stage is reached	40% Wawer et al (2005) We could find no data for MSM specifically.
Treatment trajectory	10% never get treatment 31.5% partial suppression after treatment initiation 58.5% full suppression after treatment initiation Multiple sources (see text).
Time until peak of acute viremia	21 days Little (1999)
Peak viremia	6.886 (in log10 copies per ml) Little (1999)
Time from peak viremia until set point	19 days Little (1999)
Set point	4.5 (in log10 copies per ml) Little (1999)
Time from onset of set point until AIDS-related viral increase	3610 days (for a total of 3650 days = 10 years until inset of AIDS) Buchbinder et al. 200x.
Time from onset of until AIDS-related viral increase until death	730 days (2 years); total survival with HIV without treatment is thus 12 years Buchbinder et al. 200x.

slope of viral load during AIDS	.003425 log10 copies per ml per day This implies a viral load of 7.0 at death
Viral load at full suppression	1.5 (in log10 copies per ml) Definitional.
mean viral load for those partially suppressed	3.5 (in log10 copies per ml) Chu et al. 2010
Time until partial suppression escape	365*13 days from infection Chu et al. 2010
partial.escape.slope	.003425 log10 copies per ml per day Chu et al. 2010
treatment initiation	Swindells et al. 2002. see text
Prevalence of full suppression among those on treatment	For suppression, we have two very highly concordant estimates. In HOPS, 74% of men who have consistently been on HAART for at least six months show full suppression. For Chu et al. (2010) the figure is 75%.

FIGURE 1



Source: CDC National Vital Statistics Reports Volume 56, Number 9 and National Vital Statistics Reports Volume 58, Number 8. See Table 1.

FIGURE 2. HIV prevalence over the fifty years of the simulation, for ten runs.

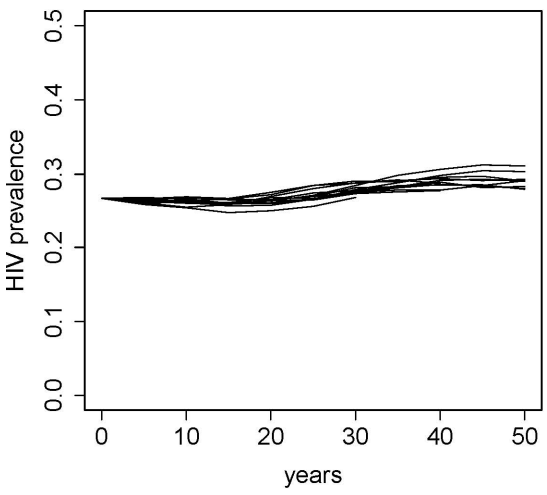


FIGURE 3: Distribution of transmission events by five variables

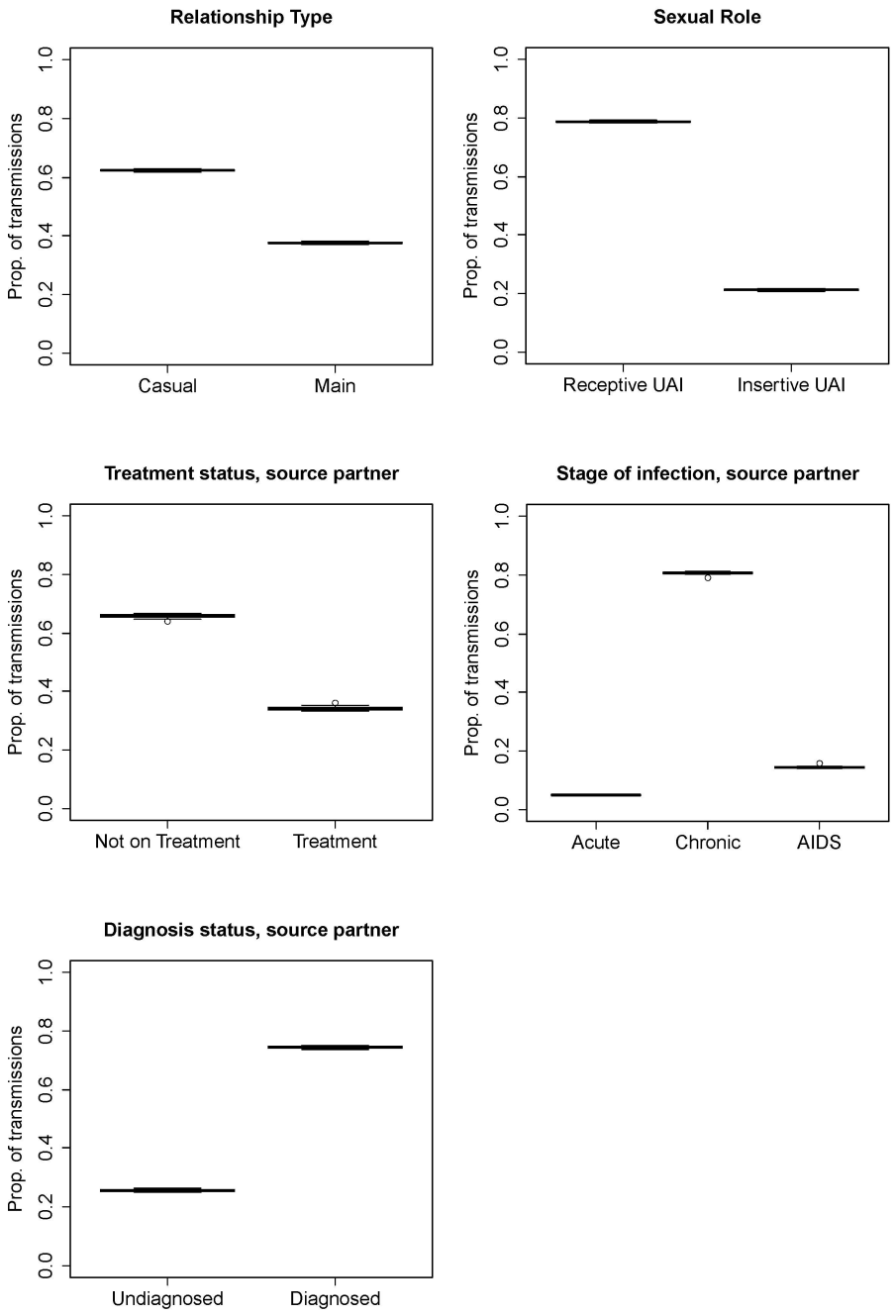


FIGURE 4

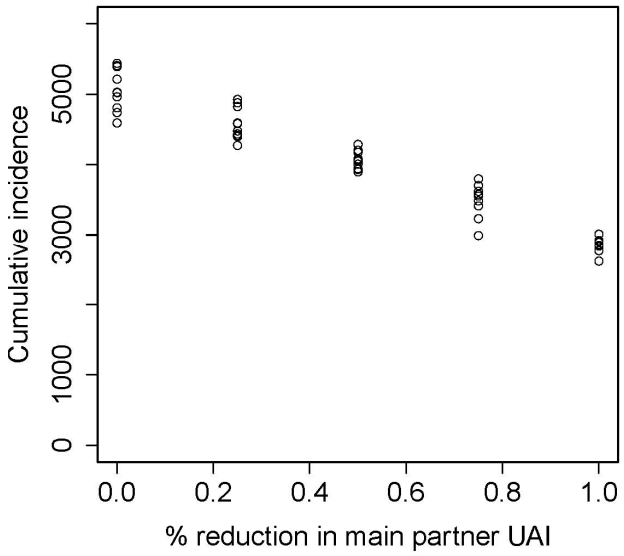

Cumulative incidence
0
1000
3000
5000
0.0
0.2
0.4
0.6
0.8
1.0
% reduction in main partner UAI

FIGURE 5. HIV prevalence by race/ethnicity