EvoNet Technical Summary

1. Model overview

Simulations were conducted using *EvoNetHIV*, a stochastic, agent-based simulation model that incorporates sexual network structure, behavior, and HIV evolution. Each simulation first estimates a statistical model that governs sexual network structure, and then proceeds through a burn-in period and epidemic simulation. At each time step of both the burn-in period and epidemic simulation, (1) partnerships form and dissolve; (2) sexual acts take place within a subset of existing partnerships; (3) HIV transmission occurs probabilistically within a subset of sexual acts; (4) viral dynamics and disease progression are updated for each infected agent; and (5) vital dynamics, such as aging, are updated. In addition, HIV treatment and preventive interventions are implemented at user-specified intervals. Each of the aforementioned processes is described in further detail below.

EvoNetHIV is programmed in the R software language (R Development Core Team 2008). Model code is accessible at https://github.com/EvoNetHIV. *EvoNetHIV* is written as a series of modules, with multiple options for each module and the option to write additional modules. It also includes over 100 parameters that users can alter, while providing default values for all of those parameters. Here we describe the *EvoNetHIV* components and parameters used in this paper; for more details, see https://github.com/EvoNetHIV/EvoNetHIV-Overview.

Simulations were conducted on the Hyak supercomputer system at University of Washington, an advanced computational, storage, and networking infrastructure provided by funding through the Student Technology Fee and the Center for Studies in Demography and Ecology.

2. Sexual network

We estimate the network using separable temporal exponential random graph models (STERGMs) (Krivitsky and Handcock 2014), as implemented in the statnet (Handcock et al. 2003) and EpiModel (Jenness et al. 2016) software suites. These algorithms also allow us to simulate a dynamic network that maintains our desired network features stochastically, even as the number of agents in the network changes, as do their attributes. Network parameterization is flexible and can be parameterized for MSM and heterosexual dynamics as well as for additional age and group structure.

3. Sexual behaviors and agent attributes

Sexual acts are determined among agents in a serodiscordant relationship at each time step. Among these partnerships, the number of sexual acts per partnership at a given time step is assigned according to a Poisson draw with mean that varies across runs (see Table 1 in manuscript body). Condom use is determined for each sexual act with probability of 50%. Circumcision status is assigned to agents at model entry with 85% probability. For MSM models, intra-event versatility (when men switch roles and each engage in insertive and receptive

anal intercourse (AI) during the same sexual contact) occurs with 40% probability when both partners are role versatile, averaged from two studies (Goodreau et al. 2012; Goodreau et al. 2017).

4. HIV transmission

The probability of transmission is calculated for each sexual act that occurs in a serodiscordant relationship, as:

$$P(transmission) = 1 - (1 - \lambda)^{e^{X\beta}}$$

where

$$X\beta = \ln(2.89) * (viral\ load - 4.0) + \ln(2.9) * insertive + \ln(17.3) * receptive + \ln(0.53) * circumcised + \ln(0.22) * condom$$

5. Set point viral load

Set point viral load (SPVL) in infected agents at model initialization is generated as a combination of viral and environmental factors. The viral contribution to SPVL is drawn from a normal distribution with mean 4.5 $\log_{10}$ copies/mL and standard deviation of $\sqrt{h^2 \times variance\ of\ log_{10}\ SPVL}$. The environmental contribution is drawn from a normal distribution with mean of 0 and standard deviation of $\sqrt{(1 - h^2) \times variance\ of\ log_{10}\ SPVL}$. SPVL is then the sum of the viral and environmental contributions, constrained to a minimum value of 2 $\log_{10}$ copies/mL and a maximum value of 7 $\log_{10}$ copies/mL.

Upon transmission, the SPVL of a newly infected agent is determined by the SPVL of the donor virus, viral mutational variance, and an environmental contribution. The viral mutational variance is drawn from a normal distribution with mean 0 and standard deviation 0.01. The environmental contribution to the SPVL of newly infected agents is drawn from the same distribution as that of infected agents at model initialization. The SPVL of newly infected agents is then the sum of the inherited SPVL of the donor agent, mutational variance, and an environmental contribution.

Table 5.1. Model parameters utilized in the assignment of set point viral load

Model parameter	**Value**	**Source(s) and notes**
Mean $\log_{10}$ SPVL at model initialization	4.5	(Fraser et al. 2007); (Korenromp et al. 2009)
Heritability of SPVL across transmissions (h^2)	0.36 (default), 0, 5	Default from (Fraser et al. 2014). Other values are sensitivity analysis.
Variance of $\log_{10}$ SPVL	0.8	(Herbeck et al. 2012)
Mutational variance	0.01	There are no published estimates of mutational variance. We have therefore programmed a low value to be conservative and to maintain approximately 0.36 heritability output measure.

6. Viral dynamics

Upon infection, viral load, *V*, grows exponentially at rate r_0 for the first 21 days according to the formula

$$V(t) = V_0 e^{r_0 t}$$

where V_0 is the initial value (set to 0.0001 copies/mL) and *t* indicates the number of days since initial infection. One study (Robb et al. 2016) has shown that viral loads during primary infection correlate with SPVL. Thus, we allowed the peak viral load to depend on the agent's SPVL as follows

$$V_{peak} = 4.639 + 0.495 * log_{10}(SPVL)$$

where the values of 4.639 and 0.495 are based on regression data given there (Robb et al. 2016). We set $r_0 = ln(V_{peak}/V_0)/21$ in order to obtain peak viral load on day 21. After reaching peak viral load, viral load decays bi-phasically. The first phase has a duration of 11 days, in which viral load decays linearly according to the following formula:

$$V(t) = V_{peak}\left(\frac{V_{32}}{V_{peak}}\right)^{\frac{(t-21)}{11}}$$

where viral load at t=32 is a weighted geometric mean of V_{adj_peak} and SPVL:

$$V_{32} = SPVL^{0.714} * V_{peak}^{0.286}$$

For the remainder of the duration of acute infection, viral load declines linearly until reaching the agent's SPVL at day 90 of infection. Viral load decay in this phase is calculated by

$$V(t) = V_{32}\left(\frac{SPVL}{V_{32}}\right)^{\frac{(t-32)}{58}}$$

In the chronic phase of HIV infection, an agent's viral load increases at a constant annual rate of 0.14 $\log_e$ copies/mL, calculated as follows

$$V(t) = SPVL * e^{0.14 * \frac{t-90}{365}}$$

This trajectory continues until an agent initiates antiretroviral treatment or enters the AIDS stage, defined by CD4 less than 200 cells/mm^3. During the AIDS stage, the agent's viral load increases linearly by 1.004112-fold per day:

$$V(t) = 1.004112 * V(t-1)$$

Viral load in AIDS increases up to a maximum viral load of 6.38 $\log_{10}$ copies/mL.

Table 6.1. Model parameters utilized in viral load dynamics

Model parameter	**Value**	**Source(s) and notes**
Viral load at day 0 of infection	0.0001	Model-calibrated to replicate viral dynamics in (Lindback et al. 2000)
r_0	1.19367006	Model-calibrated to replicate viral dynamics in (Lindback et al. 2000)
Duration of exponential viral growth	21 days	(Lindback et al. 2000)
Duration of phase 1 decay	11 days	(Lindback et al. 2000)
Duration of phase 2 decay	58 days	(Lindback et al. 2000)
Duration of acute infection	90 days	(Fiebig et al. 2003)
Viral load progression rate, natural log	0.14	(Geskus et al. 2007)
Maximum viral load in AIDS (CD4<200)	$2.4x10^6$ copies/mL = 6.38 $\log_{10}$ copies / mL	(Piatak et al. 1993)

7. Disease progression

CD4 values determine the additional risk of death among infected agents. Values are categorized as CD4 ≥ 500 cells/mm^3, 500 < CD4 ≤ 350, 350 < CD4 ≤ 200, and CD4 < 200. Agents are assigned a CD4 category probabilistically according to their set point viral load ((Cori et al. 2015); Table 7.1). No agents are assigned a CD4 category of less than 200 cells/mm^3 upon initial infection.

Table 7.1. Probability of assignment to CD4 category stratified by set point viral load

Set point viral load ($\log_{10}$ copies/mL)	CD4 level (cells/mm^3)		
	≥ 500	350 – 500	200 – 350
[2.0, 3.0]	0.88	0.12	0.00
(3.0, 3.5]	0.87	0.12	0.01
(3.5, 4.0]	0.85	0.12	0.03
(4.0, 4.5]	0.78	0.19	0.03
(4.5, 5.0]	0.73	0.21	0.05
(5.0, 5.5]	0.71	0.25	0.04
(5.5, 6.0]	0.64	0.27	0.09
(6.0, 6.5]	0.00	0.00	1.00
(6.5, 7.0]	0.00	0.00	1.00

We note that the two highest categories are included for the sake of completion, so that any individual who does evolve into this zone will have an associated CD4 value. However, these

persons are very rare in the model and die quickly, limiting the persistence of their viral genotype in the population.

In the absence of antiretroviral treatment, infected agents progress through CD4 categories probabilistically according to a geometric distribution with mean p^{-1}, where p is the inverse of the mean amount of time that an individual remains in a specified CD4 category. The mean duration of time in each CD4 category is determined by SPVL ((Cori et al. 2015) and personal communication; Table 7.2).

Table 7.2. Mean time (in years) spent in each CD4 category stratified by set point viral load

Set point viral load (log_{10} copies/mL)	CD4 level (cells/mm^3)			
	≥ 500	350 – 500	200 – 350	< 200
[2.0, 3.0]	6.08	5.01	3.60	4.67
(3.0, 3.5]	4.69	2.52	3.68	4.11
(3.5, 4.0]	3.94	4.07	2.38	3.54
(4.0, 4.5]	2.96	3.09	3.81	2.98
(4.5, 5.0]	2.25	2.32	3.21	2.42
(5.0, 5.5]	1.47	1.55	2.27	1.86
(5.5, 6.0]	0.95	1.19	1.00	1.29
(6.0, 6.5]	0.32	0.59	0.68	0.73
(6.5, 7.0]	0.30	0.46	0.37	0.17

8. Vital dynamics

8.1 Model initialization

The epidemic model can be initialized with variable population sizes, though a maximum of 5,000 is considered the upper limit due to computational resources required for larger sizes. The default initial age distribution of model agents is obtained for United States males ages 18-85 from Centers for Disease Control and Prevention (CDC) Wide-ranging Online Data for Epidemiologic Research (WONDER) data for the years 1999-2003 (Centers for Disease Control and Prevention 2015). This age distribution was used in a model without treatment, reflecting the high AIDS mortality rate observed in the first two decades of the AIDS epidemic, until reaching an equilibrium with respect to age. This equilibrium age distribution is scaled to the age range of 18-55, such that the sum of proportions of agents in each age category is equal to 1. The age of each agent is then randomly assigned with probability of a given age equal to the proportion of the scaled equilibrium U.S. male population of that age.

8.2 Entries

The number of entries (births) into the model at each time step is determined by a Poisson draw from a distribution with mean 1.37. This distribution results in approximately 1% annual population growth when all of the default Evonet parameters are used. Each new agent enters the model uninfected with age 18.

8.3 Exits

Age-specific annual mortality rates for US males ages 18-55 were obtained from the CDC WONDER database for the years 1999-2003 (Centers for Disease Control and Prevention 2015). We converted these annual mortality rates to daily probabilities.

Natural deaths occur according to each agent's age-specific probability of death, and are determined probabilistically by a random draw from a uniform distribution on [0, 1]. HIV-infected agents with CD4 greater than 200 cells/mm^3 have an increased probability of death that is dependent on their CD4 category.

Deaths due to AIDS occur when an agent's time in CD4 category 4 (CD4 < 200 cells/mm^3) is completed according to disease progression described in Section 7.

8.4 Aging

Each agent's age is incremented by 1/365 at each time step.

Table 8.1. Model parameters governing vital dynamics

Model parameter	**Value**	**Source(s) and notes**
Initial population size	5,000	NA
λ for model entries (births)	1.37	Model-calibrated to produce 1% annual growth
Minimum age	18	NA
Maximum age	55	NA
Age distribution	0.0450, 0.0440, 0.0430, 0.0420, 0.0410, 0.0400, 0.0390, 0.0380, 0.0370, 0.0360, 0.0350, 0.0340, 0.0330, 0.0320, 0.0310, 0.0300, 0.0290, 0.0280, 0.0270, 0.0260, 0.0250, 0.0240, 0.0230, 0.0220, 0.0210, 0.0200, 0.0190, 0.0180, 0.0170, 0.0160, 0.0150, 0.0140, 0.0130, 0.0120, 0.0110, 0.0100, 0.0090	Modified from CDC WONDER (Centers for Disease Control and Prevention 2015)
Age-specific annual mortality rates	0.0011, 0.0012, 0.0013, 0.0014, 0.0014, 0.0014, 0.0014, 0.0014, 0.0014, 0.0014, 0.0014, 0.0014, 0.0015, 0.0015, 0.0016, 0.0016, 0.0017, 0.0018, 0.0019, 0.0021, 0.0022, 0.0024, 0.0026, 0.0028, 0.0030, 0.0033, 0.0036, 0.0039, 0.0043, 0.0046, 0.0050, 0.0055, 0.0059, 0.0064, 0.0069, 0.0074	CDC WONDER (Centers for Disease Control and Prevention 2015)
Additional probability of death with CD4 > 500 cells/mm^3	0.0000112 per day	The values in CASCADE, 2011 (Writing Committee for the CASCADE Collaboration 2011)

		are for men with mean age 30. Rates presented here therefore subtract 0.0014, the natural mortality rate for North American males aged 30 (Centers for Disease Control and Prevention 2015), to estimate an excess death rate associated with this CD4 category.
Additional probability of death with CD4 350-500 cells/mm^3	0.0000148 per day	See note above
Additional probability of death with CD4 200-350 cells/mm^3	0.0000333 per day	See note above

References

Centers for Disease Control and Prevention (2015), *Wide-ranging Online Data for Epidemiologic Research*. (2015).

Cori, A., et al. (2015), 'CD4+ cell dynamics in untreated HIV-1 infection: overall rates, and effects of age, viral load, sex and calendar time', *AIDS,* 29 (18), 2435-46.

Fiebig, E. W., et al. (2003), 'Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection', *Aids,* 17 (13), 1871-79.

Fraser, C., et al. (2007), 'Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis', *Proc Natl Acad Sci U S A,* 104 (44), 17441-6.

Fraser, C., et al. (2014), 'Virulence and pathogenesis of HIV-1 infection: an evolutionary perspective', *Science,* 343 (6177), 1243727.

Geskus, R. B., et al. (2007), 'The HIV RNA setpoint theory revisited', *Retrovirology,* 4.

Goodreau, S. M., et al. (2017), 'Sources of racial disparities in HIV prevalence in men who have sex with men in Atlanta, GA, USA: a modelling study', *Lancet HIV,* 4 (7), e311-e20.

Goodreau, S. M., et al. (2012), 'What drives the US and Peruvian HIV epidemics in men who have sex with men (MSM)?', *Plos One,* 7 (11), e50522.

Handcock, M. S., et al. (2003), 'statnet: Software tools for the Statistical Modeling of Network Data'.

Herbeck, J. T., et al. (2012), 'Is the virulence of HIV changing? A meta-analysis of trends in prognostic markers of HIV disease progression and transmission', *AIDS,* 26 (2), 193-205.

Jenness, S.M. , Goodreau, S.M. , and Morris, M. (2016), 'EpiModel: Mathematical Modeling of Infectious Disease', *R Package Version 1.2.7. URL: http://epimodel.org/. DOI: 10.5281/zenodo.16767.*

Korenromp, E. L., et al. (2009), 'Clinical Prognostic Value of RNA Viral Load and CD4 Cell Counts during Untreated HIV-1 Infection-A Quantitative Review', *Plos One,* 4 (6).

Krivitsky, P. N. and Handcock, M. S. (2014), 'A separable model for dynamic networks', *Journal of the Royal Statistical Society Series B-Statistical Methodology,* 76 (1), 29-46.

Lindback, S., et al. (2000), 'Viral dynamics in primary HIV-1 infection', *Aids,* 14 (15), 2283-91.

Piatak, M., et al. (1993), 'High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR', *Science,* 259 (5102), 1749-54.

R Development Core Team (2008), 'R: A language and environment for statistical computing'.

Robb, M. L., et al. (2016), 'Prospective Study of Acute HIV-1 Infection in Adults in East Africa and Thailand', *New England Journal of Medicine,* 374 (22), 2120-30.

Writing Committee for the CASCADE Collaboration (2011), 'Timing of HAART Initiation and Clinical Outcomes in Human Immunodeficiency Virus Type 1 Seroconverters', *Archives of Internal Medicine,* 171 (17), 1560-69.