Bayesian Estimate of the Parameters of a Stochastic Differential Model of HIV Incidence in the United States

Renee Dale[1,2] & BeiBei Guo[1]

[1]Department of Experimental Statistics
[2]Department of Biological Sciences
Louisiana State University

Human Immuno-deficiency Virus

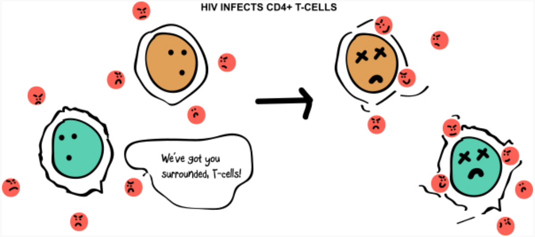
The human immunodeficiency virus (HIV) progresses in three stages:

- The first stage lasts approximately 3 months
- The chronic stage can last from 5-10 years without medication
- The disease then progresses to auto-immunodeficiency syndrome (AIDS)

An infected individual may go a long time before they are diagnosed while still contributing to the epidemic.

Estimation of Undiagnosed Prevalence

CD4 cell level is used as an indicator of the progression of the disease. As the disease progresses the CD4 levels drop.

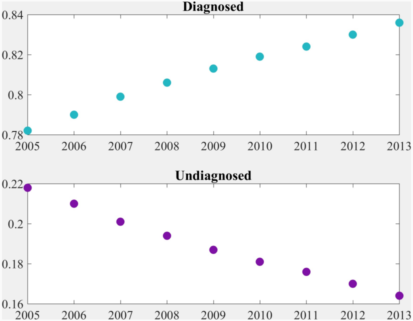


1

[1]Song, R., et al. Using CD4 Data to Estimate HIV Incidence, Prevalence, and Percent of Undiagnosed Infections in the United States. JAIDS 74, 3-9 (2017)

Problems with models

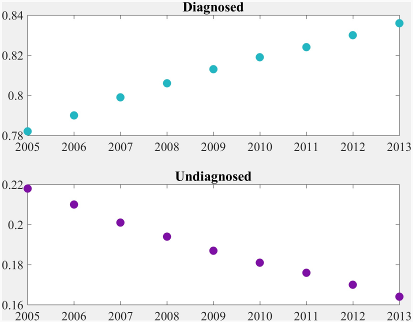
The biggest problem with dynamic models is specifying the parameters. For the case of HIV dynamics, the transmission and diagnosis rate are the key parameters.



Since some percent of the population is unknowable, it is very difficult to estimate these parameters.

Objectives

- Derive epidemiological parameters in a new way
- Permutate these parameters to understand dynamics



Bayesian Model

Use Bayesian statistics to find the *proportional change* for each of the HIV infected populations:

$$p_t = \mathbf{q} p_{t-1}$$

First we identify a sampling model that represents our data.

A **binomial distribution** makes sense for looking at the chance of finding an undiagnosed or diagnosed individual in the population of HIV infected individuals:

$$Bin(x_t; n_t, p_t)$$

Bayesian Model

The prior for **q** is:

$$\pi(q) \propto Gamma(\alpha, \beta)$$

centered at the arithmetic estimate from prior literature: 0.979 for undiagnosed and 1.025 for diagnosed.
The prior for the random variable p_t is a beta centered at the previous proportion.

$$\pi(p_t) \propto P(\mathbf{q} p_{t-1} n_t - 1, n_t - \alpha)$$

In the case where t=1, the previous proportion is taken to be the expert opinion of 20%.

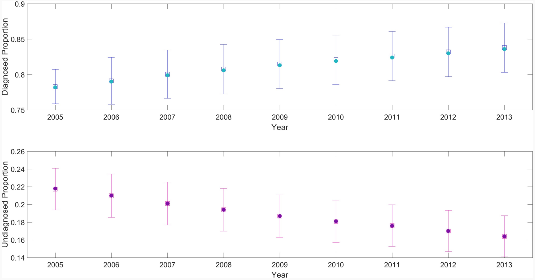
Bayesian Model

The joint posterior distribution is proportional to the priors multiplied by the likelihoods for all 9 years:

$$\Pr(p_1, p_2, ..., p_9, q) \propto \pi(q) \times \Pi_{t=1}^{9} \pi(p_t | q, p_{t-1})$$
$$\times \mathcal{L}(p_1 | x_{t=1}) \times \mathcal{L}(p_2 | x_{t=2}) \times ... \times \mathcal{L}(p_9 | x_{t=9})$$

Since the posterior distribution doesn't have a closed form, Metropolis-Hastings nested within a Gibbs sampler is used to sample from the posterior for 100,000 iterations.

Results



Parameter	Posterior Mean
q_u	0.978
q_d	1.036

Stochastic Differential Equations

We can use the proportional change to estimate important epidemiological parameters transmission and diagnosis.

The proportional change estimates are used as constraints on the process:

$$dU = (\mathbf{q_u} - 1)Udt + d\omega_t dt$$

$$dD = (\mathbf{q_d} - 1)Ddt + d\omega_t dt$$

where $d\omega_t \sim Nor(0, \sigma)$ is Brownian white noise.

Epidemiological Parameter Estimation

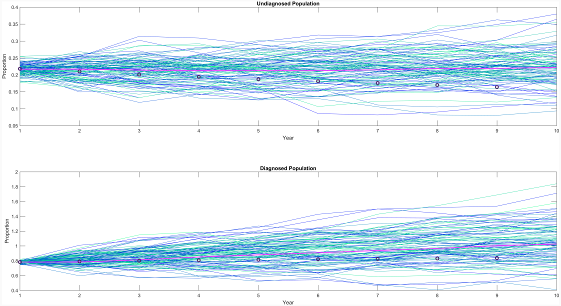
The simplest model of this process has 3 parameters: **transmission** (τ), **diagnosis** (δ), and **death** (ϵ).

$$dU = (\boldsymbol{\tau}(U + D) - \boldsymbol{\delta} U - \boldsymbol{\epsilon} U)dt \cong (1 - \mathbf{q_u})Udt + d\omega_t dt$$

$$dD = (\boldsymbol{\delta} U - \boldsymbol{\epsilon} D)dt \cong (1 - \mathbf{q_d})Ddt + d\omega_t dt$$

Using the estimates for q_d and q_u, we can get estimates for these parameters.

Base Model



$$dU = (\tau(U + D) - \delta U - \epsilon U)dt \cong (1 - \mathbf{q_u})Udt + d\omega_t dt$$

$$dD = (\delta U - \epsilon D)dt \cong (1 - \mathbf{q_d})Ddt + d\omega_t dt$$

$$\tau = 0.0334 \qquad \delta = 0.165$$

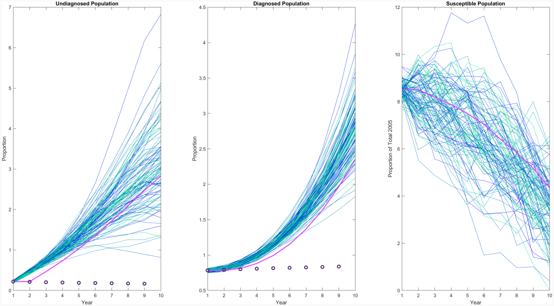
Understanding the population dynamics

Exhaustion of Susceptibles
Running out of high-risk susceptible individuals would cause the undiagnosed population to decrease over time.

Lack of Access to Care
Some geographic and socio-economic groups have higher rates of HIV and lack consistent access to care.

Anti Retroviral Therapy
96% of infected individuals reported being on ART in a recent study.

Susceptible Exhaustion

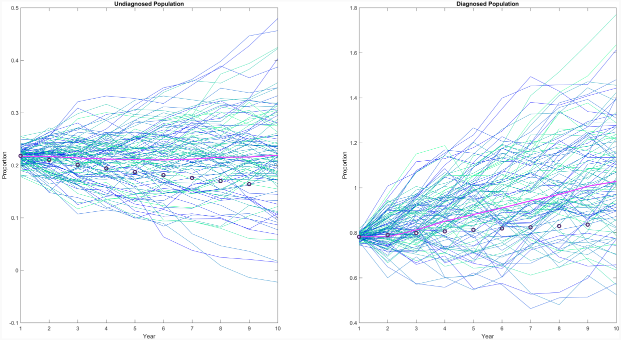


For $S \cong T$, transmission is a function of S: $S = fT$

The transmission term becomes

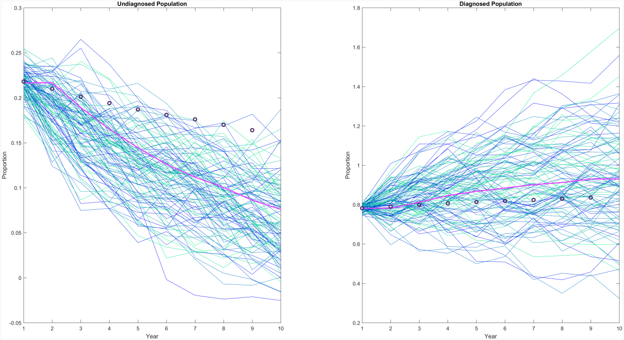
$$\tau TS \cong \tau f T^2$$

Lack of Access to Care



For this case, we consider diagnosis as a **constant** that does not depend on the size of the undiagnosed population.

ART Usage



The size of the infected population able to transmit the disease is reduced by $0.96D$.

Parameter sets

Model	**Transmission Rate**	**Diagnosis Rate**
Base model	$\tau(U+D)$	δU
Exhaustion of Susceptibles (ES)	$\tau f(U+D)^2$	δU
Lack of Access to Care (LAC)	$\tau(U+D)$	δ_0
Anti-retroviral Therapies (ART)	$\tau(U+0.04D)$	δU
ES and LAC	$\tau f(U+D)^2$	δ_0
ES and ART	$\tau f(U+0.04D)^2$	δU
LAC and ART	$\tau(U+0.04D)$	δ_0
ES, LAC, and ART	$\tau f(U+0.04D)^2$	δ_0

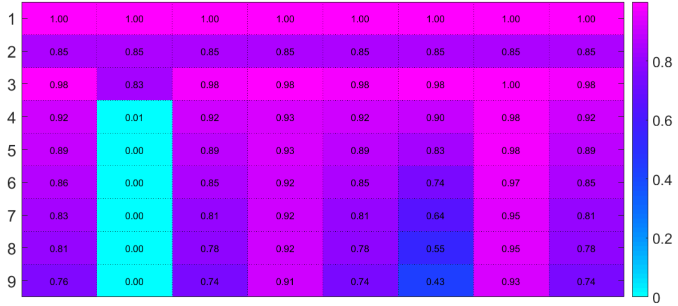
Conclusion

Undiagnosed

Year	H_0	H_1	H_2	H_3	$H_1 H_2$	$H_1 H_3$	$H_2 H_3$	$H_1 H_2 H_3$
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91
3	0.89	0.00	0.90	0.91	0.00	0.25	0.69	0.29
4	0.88	0.00	0.91	0.73	0.00	0.10	0.63	0.13
5	0.79	0.00	0.87	0.49	0.00	0.03	0.49	0.05
6	0.74	0.00	0.86	0.36	0.00	0.02	0.44	0.04
7	0.69	0.00	0.85	0.28	0.00	0.03	0.41	0.04
8	0.58	0.00	0.81	0.23	0.00	0.04	0.37	0.05
9	0.53	0.00	0.80	0.17	0.00	0.08	0.35	0.08
Avg P	0.7784	0.2121	**0.8799**	**0.5640**	0.2120	0.2742	**0.5873**	0.2877

Conclusion

Diagnosed



Year	H_0	H_1	H_2	H_3	$H_1 H_2$	$H_1 H_3$	$H_2 H_3$	$H_1 H_2 H_3$
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
3	0.98	0.83	0.98	0.98	0.98	0.98	1.00	0.98
4	0.92	0.01	0.92	0.93	0.92	0.90	0.98	0.92
5	0.89	0.00	0.89	0.93	0.89	0.83	0.98	0.89
6	0.86	0.00	0.85	0.92	0.85	0.74	0.97	0.85
7	0.83	0.00	0.81	0.92	0.81	0.64	0.95	0.81
8	0.81	0.00	0.78	0.92	0.78	0.55	0.95	0.78
9	0.76	0.00	0.74	0.91	0.74	0.43	0.93	0.74
Avg P	0.8790	0.2989	**0.8693**	**0.9302**	0.8693	0.7705	**0.9558**	**0.8693**

Conclusion

These results suggest:

- The proportional change of a population can be used to constrain parameters of a dynamical model
- The population dynamics of the diagnosed and undiagnosed populations are best explained by a lack of access to care and ART usage
- The diagnosis rate is estimated to be between 3.4% and 16%, lower than the previously estimated 20%
- The transmission rate is estimated to be 3.3% and depend on only part of the diagnosed population. This is close to the previously estimated 4%

Thank you!

This research was partially supported by the NIH R01 grant AI12125903