EpiModel

An R Package for Mathematical Modeling of Infectious Diseases over Networks

UseR! 2017 Conference
Brussels, Belgium | July 5, 2017

Samuel M. Jenness, PhD MPH
EpiModel Research Lab
Department of Epidemiology // Rollins SPH // Emory University

samueljenness.org  smjenness  @SamuelJenness
What is EpiModel?

• An R package!
• Part of the Statnet suite
• Open-source software platform for epidemic modeling
• Tools for building, simulating, and analyzing models
• Three model classes
  - Deterministic compartmental models
  - Stochastic ABMs
  - Stochastic network models
• http://epimodel.org/
Person 1 has non-zero risk of infection as a result of his partner’s risk.

Person 2 exhibits same individual-level behavior, but not (currently) at risk.
Why Networks Matter for HIV/STI Epidemiology


Small changes in individual behavior result in large changes to the network connectivity

Mathematical Models for Infectious Disease Dynamics

In contrast to empirical study designs, mathematical models:

- **Abstract** the processes of ID dynamics to understand the data generating mechanism
- **Project up** to the population level
- **Project forward** over time
- Handle complex contagion dependencies

Simple SIR Model (e.g., Measles)

Contagion = Dependencies

Population-Level Outcomes over Time
1. Each new population heterogeneity **multiplies** number of equations to solve
   - Realistic forms of population heterogeneity could involve hundreds to thousands of ODEs

2. DCMs can implement only a limited form of repeated contacts within partnerships
   - The evaluation of the individual acts is compressed to the start of those partnerships

---

**The Reality**

- 10%  •  •  •  10%  •  •  •  10%

- $t_1$  •  •  •  $t_3$

---

**The Math**

\[
\begin{align*}
\frac{d(X_i)}{dt} &= \mu T_i + \alpha (A_i^{0,0} + A_i^{1,0} + (1 - \varepsilon_M)A_i^{1,1}) - \lambda_i - \mu X_i \\
\frac{d(U_i)}{dt} &= \lambda_i - (\nu + \mu) U_i \\
\frac{d(E_i^0)}{dt} &= \nu U_i - (\rho + \tau_i + \mu) E_i^0 \\
\frac{d(E_i^1)}{dt} &= \tau_i E_i^0 - (\rho + \mu) E_i^1 \\
\frac{d(Y_i^0)}{dt} &= \rho E_i^0 - (\theta_i + \kappa + \mu) Y_i^0 \\
\frac{d(Y_i^1)}{dt} &= \rho E_i^1 + \theta_i Y_i^0 - (\kappa + \mu) Y_i^1 \\
\frac{d(Z_i^{0,0})}{dt} &= \kappa Y_i^0 - (\theta_i + \gamma + \mu) Z_i^{0,0} \\
\frac{d(Z_i^{1,0})}{dt} &= \kappa Y_i^1 + \theta_i Z_i^{0,0} + \omega Z_i^{1,1} - (\gamma + \mu) Z_i^{1,0} \\
\frac{d(Z_i^{1,1})}{dt} &= Z_i^{1,0} - (1 - \varepsilon_C) E_i^{1,1}
\end{align*}
\]
Definition of Network Models

- Network models explicitly represent **repeated contacts between the same people** over time
- That set of contacts may stay **static** over the epi dynamics
  - Acute immunizing infections
- Or potentially **changing** across time steps, along with the epi dynamics
  - Long-term non-immunizing infections
- Most network models are **agent-based** or **individual-based**
**Graph probability**

\[ P(Y = y) = \frac{\exp\{\theta' g(y)\}}{\mathcal{K}(\theta)} \]

- \( g(y) \) = vector of network statistics
- \( \theta \) = vector of parameters
- \( \mathcal{K} \) = normalizing constant

- Probability of network \( Y \) is an exponential function of network statistics with associated parameter values
- Can be re-expressed as a conditional logit for dyads \( Y_{ij} \)
- Because of denominator (all possible networks) simulation-based MCMC methods used for estimation
Can We Observe the Y (the Network) for HIV?

- Phylogenetic networks
  - Probabilistic transmission clusters
  - Transmission directionality cannot be inferred without external (clinical) data
  - Answering a different (but still important) set of questions

- Network “census”
  - Likoma Island (Malawi) partnership network measured through survey data
  - Around 50% missing data
  - Not dynamic
Network Sampling Methods

Census

- Data on every node and every link
- Ideal, but unlikely
- Missing data on edges very problematic

Adaptive Sampling

- Variations on link tracing, snowball sampling, contact tracing
- Work on statistical inference in this area, but many assumptions involved

Egocentric Sampling

- Sample persons, ask them to report on partners
- Use a statistical model to estimate population parameters from sample
- With a properly specified model, simulate!
Example $g(y)$ Network Parameters

An Arbitrarily Complex Set of Interactions
• TERGM = temporal ERGMs
• Separate ERGMs for formation and dissolution
• Dissolution parameters also estimated by cross-sectional data
A suite of R packages for analysis, visualization, statistical modeling, (and now) epidemic modeling with networks

http://statnet.org/

https://CRAN.R-project.org/package=statnet
• Dynamic network simulation placed on top of an agent-based model (ABM) for an infectious disease

• ABM defines the possible states a person may be in, and transitions between states over time

• New students start with EpiModel’s “base” models: arbitrarily network configurations + fixed ABM for epi
• Standard templates may be expanded to address novel research questions
• Possible with a standardized Application Programming Interface (API) for building new structure into the model
• Modules to be added or modified perform one unique process within the system

```cpp
EpiModel::netsim(est, param, init, control)
    t1 initialization module
    for (t in 2 to t_n) {
        module 1
        module 2
        network re-simulation module
        transmission module
        module 5
    }
    t_n Clean up and save output
```
Building Modules for HIV/STI Research

- Diagnosis
- Treatment
- Viral Load
- STI Co-Inf.
- CD4 Count
- Transmission
- Mortality
**Problem:** How does sexual partnership network structure improve or reduce efficiency of proven interventions for HIV (MMC)?

**Model:** Model parameterized with network data collected in Ghana; sensitivity analyses varied sex-specific concurrency.

**Results:** Increase connectivity in partnership networks results in decreased efficiency but overall HIV infections averted of biomedical circumcision interventions.

Application 2: Modeling HIV Preexposure Prophylaxis

• **Problem:** How to most efficiently allocate HIV preexposure prophylaxis (PrEP) to men who have sex with men in the US?

• **Model:** CDC PrEP indicated for men based on complex partner network criteria; many uncertainties in PrEP uptake and adherence

• **Results:** HIV PrEP could prevent 33% of new HIV infections over the next decade if clinicians follow CDC’s guidelines

EpiModel Resources

Github

github.com/statnet/EpiModel

Methods Paper in Press @ JSS

Pre-press Version on Website

EpiModel: A R Package for Mathematical Modeling of Infectious Disease over Networks

Samuel M. Jemness
Emory University

Steven M. Goodreau
University of Washington

Martina Morris
University of Washington

Abstract

EpiModel provides tools for building, numerically solving, and plotting mathematical models for the population dynamics of infectious disease transmission in R. Many classes of models are included, but the unique feature of this software is a general stochastic framework for modeling the spread of epidemics on networks. This framework integrates recent advances in statistical methods for network analysis, temporal exponential random graph models, and allows the epidemic modeling to be firmly grounded in empirical data on the contacts and partnerships that can spread infection. This article provides an overview of both the general modeling tools included in EpiModel, which are designed to facilitate teaching and learning for students new to modeling, and the application programming interface, which is designed to facilitate developing extensions to EpiModel to address novel research questions.

Keywords: mathematical model, infectious disease, epidemiology, networks, R

1. Introduction

The EpiModel package provides tools for building, simulating, and analyzing mathematical...
Future Directions for EpiModel

1. **Extending** other common forms of network data or information relevant for HIV/STI and other infectious diseases

2. **Developing** further complexity of HIV/STI modules to address critical questions in prevention and control

3. **Integrating** health economic analyses (CEA) within model outputs for public health policymaking

4. **Supporting** students, researchers, policymakers and other potential EpiModel users
Acknowledgements

Network Modeling Group @ UW & Co-Authors
Martina Morris
Steve Goodreau

Statnet Development Group
Mark Hancock
Pavel Krivitsky
Dave Hunter
Carter Butts
Skye Bender deMoll

EpiModel Research Lab @ Emory
Kevin Weiss
Emeli Anderson
Jeb Jones
Maraia Tremarelli

NIH R01HD68395: Statistical Methods for Network Epidemiology (PI: Martina Morris)

NIH R21HD075662: Using Sexual Network Transmission Models to Explain HIV Disparities Between Black and White MSM (PI: Steven Goodreau)

NIH R01AI108490: Integrated Bio-Social Models for HIV Epidemiology (PIs: Steven Goodreau, Joshua Herbeck, and John Mittler)

NIH R21MH112449: Modeling Antiretroviral-Based Prevention among MSM in the US (PI: Samuel Jenness)

NIH P30AI050409: Center for AIDS Research at Emory University (PIs: Carlos del Rio and James Curran)

NIH P30DA027828: Center for Prevention Implementation Methodology for Drug Abuse and HIV (Ce-PIM) (PIs: Henricks Brown and Brian Mustanski)