

# PROTOCOL FOR:

## Sex-assortativity and the spread of TB on contact networks

*Spring 2019*

### Authors:

- Paige Miller
- John Drake

### Background:

TB is a respiratory-transmitted infectious disease that is heterogeneously distributed globally and concentrated in Africa and Asia. In addition, TB is heterogeneously distributed within populations; in particular, men face higher risk factors than women. The global male:female TB case ratio is 1.9, a pattern strikingly consistent across countries.

Male-bias in TB case reports is thought to be due to sex-specific differences in exposure to Mtb or susceptibility to disease following exposure, or a combination of the two. Exposure rates could be mediated by differences in behaviors and societal gender roles. For example men may have more contacts or be more central in social networks though previous studies have failed to substantiate this (Mosson et al. 2008, etc.). Alternatively, sex-specific exposure rates could be driven by differences in the types of contacts that men compared to women. Specifically, exposure could be higher in men due to assortative mixing which is where individuals tend to associate with others similar to them, creating sub-groups within a social network. However, we lack studies analyzing the role of hypothesized exposure mechanisms at driving male-bias.

The effects of assortativity (i.e., modularity) on aggregated (i.e., regardless of subgroup) epidemic dynamics (e.g., total outbreak size or equilibrium prevalence) depend on the infection process (SIR, SIS) and the distribution of subgroup size. Tuberculosis transmission is better represented by an SIR than an SIS process. In SIR systems, the total outbreak size of diseases with SIR dynamics decrease with increasing assortativity, perhaps due to a build-up of recovered nodes within-subgroups and transmission bottlenecks between subgroups (Nadimi 2018). However, very high assortativity may be required to produce this “protective” effect of assortativity (Salathe & Jones 2015, Sah et al. 2017). In contrast, for SIS processes, infected nodes return to susceptible nodes and a build-up of recovered nodes does not occur, resulting in intensely connected subgroups which increase the equilibrium prevalence of infection (Nadimi 2018). Subgroup size also affects disease spread: networks with larger subgroups increase epidemic size when compared with networks with constant subgroup size (Sah et al. 2017). The number of subgroups also affects how assortativity relates to outbreak size: outbreak size on less fragmented (fewer subgroups) networks are not affected by assortativity but outbreak size on more fragmented networks increases with assortativity (Sah et al. 2017). Most prior studies have focused on networks with multiple (in the range of 10 to 1000, Sah et al. 2017), highly connected subgroups. For our purposes, an open question remains about how disease spreads within lightly assorted networks with two subgroups (i.e., men and women) which potentially vary in their susceptibility to infection (i.e., are men “supersusceptibles”, Kraft 2015).

We wondered whether sex-assortativity in social networks could contribute to male-bias or if previously described differences in susceptibility between the sexes are necessary to explain male-bias.

### Research questions:

1. Can sex-assortative mixing lead to male-bias by itself or is sex-specific susceptibility required to explain male-bias?

2. What are the effects of sex-assortative mixing on disease spread (peak size/time, variation in outbreak size/duration)?

### Study design:

We will examine the effects of sex-assortativity and sex-specific susceptibility on the ratio of male to female cases using models of disease spread on simulated networks. In the simulated networks, each node will represent an individual and each link is a connection between individuals that infection can spread.

#### Network generation:

Simulated networks will vary in level of sex-assortativity,  $r$ , calculated according to Newman's discrete assortativity coefficient (Newman 2003) as

$$r = \frac{\sum_i e_{ii} - \sum_i (a(i)^2)}{1 - \sum_i (1 - a(i)^2)} = \frac{Tr\mathbf{E} - ||E^2||}{1 - ||E^2||}$$

where  $e_{ij}$  is the proportion of edges connecting nodes in subgroup  $i$  to subgroup  $j$  (undirected),  $a(i) = \sum_j e_{ij}$ . Alternatively, if  $\mathbf{E}$  is the matrix of  $e_{ij}$  and  $||X||$  is the sum of all elements in a matrix  $\mathbf{X}$ , then assortativity can be calculated with the proportion of edges within-sex ( $Tr\mathbf{E}$ ) and the proportion of edges that would be within groups if connections were random  $||E^2||$ . Here, if edges were distributed randomly,  $Tr\mathbf{E} = ||E^2|| = 0.5$  and  $r = 0$ . Typically,  $r$  ranges from  $-1 \gg r \geq 1$  because disassortative networks are much closer to random networks ( $r=0$ ) than are assortative networks.

Note: Network modularity ( $Q$ ) is a similar measure of non-random mixing in networks:  $Q = \sum_i e_{ii} - a_i^2 = Tr\mathbf{E} - ||E^2||$  where  $e_{ij}$  is the proportion of edges in the network that link nodes in community  $i$  to community  $j$  and  $a_i = \sum_j e_{ij}$  represents the proportion of edges in the network that link to nodes in subgroup  $i$ . The maximum value of  $Q$  is  $1 - \frac{1}{K}$  where  $K$  is the number of modules (Sah 2014). Thus, the relationship between assortativity and modularity for networks with two subgroups is

$$assortativity = modularity / (1 - expected.prop.within.edges)$$

where the expected proportion of within edges is the proportion of nodes in that subgroup. Since groups have equal size, you would expect 1/2 to occur within group by chance. Thus, assortativity gets divided by 0.5 while modularity does not.

We will generate scale-free networks according to the parameters listed in Table 1 using the classic BA-algorithm. Following network generation, we will update the networks as following:

1. Assign nodes randomly as male (0) or female (1).
2. Calculate temporary value of sex-assortativity in the network ( $r_t$ ).
3. If  $r_t$  is not within  $\epsilon$  of  $r$ , randomly choose a proportion  $\alpha$  of 0–1 edges (i.e, a male–female edge) if  $r \geq 0$  and re-wire them or if  $r < 0$ , choose a proportion  $\alpha$  of 0–0 and 1–1 edges and re-wire them.
4. Repeat step 3 until  $|r_f - r_t| \leq \epsilon$ .

Network will be generated with parameters shown in Table 1.

Table 1: Network parameters.

Variable	Value
Sex-assortativity, $r$	(0, 0.9) by 0.1
Degree distribution, $p(K)$	$\frac{k^{-\alpha}}{\zeta(\alpha)}$
Mean degree, $\langle K \rangle$	10
Network size, $N$	$1 \cdot 10^3$
Tolerance, $\epsilon$	0.035

Variable	Value
Rewiring proportion, $\alpha$	0.2
Replicates	100

#### *Models of disease transmission:*

We were interested in comparing models of TB that assume outbreak (SIR) and persistent levels of infection (SIRS).

Due to sex-specific susceptibility, models will assume a non-Markovian process at first because the transmission rate will vary depending on susceptibility of each node. Once susceptibility is assigned, the process can be treated as Markovian (see Kiss, Miller, and Simon, page 224). We will use an event-driven, non-Markovian algorithm to implement simulations (Kiss, Miller & Simon 2017).

To hold mean node susceptibility to 1, with a m:f ratio of  $\alpha$ , we solved the following equations to find  $\sigma_m, \sigma_f$ :

$$\begin{aligned} 0.5\sigma_m + 0.5\sigma_f &= 1 \\ \sigma_m &= \alpha\sigma_f \end{aligned}$$

Leading to solutions:

$$\begin{aligned} \sigma_f &= \frac{2}{\alpha + 1} \\ \sigma_m &= \frac{2\alpha}{\alpha + 1} \end{aligned}$$

#### **SIR:**

An infection spreads along an edge at probability depending on the baseline transmission rate,  $\tau$ , the susceptibility of the target node, and the duration of infection, T:  $p(T) = 1 - e^{-\tau\alpha T}$ . Susceptibility will be altered by changing  $\alpha$ , the ratio of male:female susceptibility. Infecteds recover at exponentially distributed recovery rate  $\gamma$ .

Parameters for the SIR model are given in Table 2.

Table 2: Disease parameters for SIR model.

Variable	Value
Initial susceptible, $S_0$	$N - 0.05 \cdot N$
Initial infected, $I_0$	$0.05 \cdot N$
Infection rate, $\tau$	0.08, 0.12, 0.16, 0.2, .24, .28, .32, 0.36
Recovery rate, $\gamma$	1
M:F susceptibility ratio, $\alpha$	1, 1.25, 1.5, 1.75

#### **SIRS:**

The SIRS model can be interpreted as an SIR model with births (Keeling & Eames 2005).

#### **Analysis:**

##### *Network structure:*

- Relationship between network structure (clustering, degree assortativity, diameter, degree variation) and assortativity

*M:F ratio:*

- Relationship between assortativity and M:F case ratio
- Relationship between susceptibility and M:F case ratio
- Interaction between assortativity and susceptibility on M:F case ratio
- Variation by network size and  $R_0$

*Disease spread:*

- Relationship between assortativity and epidemic peak size, peak time, variation in outbreak size, variation in epidemic duration
- Variation by network size and  $R_0$

**Checklist:**

- X Understand the relationship between measures of community structure ( $Q$  vs.  $r$ ) for  $K=2$  modules
- X Run study across larger parameter grid and more replicates
- X Analyze and interpret results from extended simulations
- X Understand and relate results to Salathe and Sah research
- X Decide on next steps which could be: (1) Seed epidemics disproportionately in one module; (2) Incorporate sex-specific susceptibility; (3) Incorporate latent class of individuals; (4) Sample epidemics according to COHSONET and validate results
- X edit protocol to add variable susceptibility for SIR
- X Analyze pilot study of variable susceptibility for SIR
- Run extended analysis of variable susceptibility for SIR
- Figure out SIRS model with variable susceptibility
- Run pilot of SIRS model with variable susceptibility
- Run extended SIRS model
- Write up analysis with variable susceptibility

**Important background papers:**

Kiss, I Z, J C Miller, and PL Simon Cham Springer. 2017. “Mathematics of Epidemics on Networks.” Springer.

Miller, Joel C. 2007. “Epidemic Size and Probability in Populations with Heterogeneous Infectivity and Susceptibility.” *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* 76 (1 Pt 1): 010101. doi:10.1103/PhysRevE.76.010101.

Newman, MEJ. 2003. “Mixing Patterns in Networks.” *Physical Review E* 67 (2). American Physical Society. doi:10.1103/PhysRevE.67.026126.

Nhamoyebonde, Shepherd, and Alasdair Leslie. 2014. “Biological Differences Between the Sexes and Susceptibility to Tuberculosis.” *Journal of Infectious Diseases* 209 (suppl 3): S100–S106. doi:10.1093/infdis/jiu147.

Pastor-Satorras, R, and A Vespignani. 2001. “Epidemic Dynamics and Endemic States in Complex Networks.” *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* 63 (6 Pt 2): 066117. doi:10.1103/PhysRevE.63.066117.

Perkins, S E, M F Ferrari, and P J Hudson. 2008. “The Effects of Social Structure and Sex-Biased Transmission on Macroparasite Infection.” *Parasitology* 135 (13): 1561–69. doi:10.1017/S0031182008000449.

Sah, Pratha, Stephan T Leu, Paul C Cross, Peter J Hudson, and Shweta Bansal. 2017. “Unraveling the Disease Consequences and Mechanisms of Modular Structure in Animal Social Networks.” *Proceedings of*

the National Academy of Sciences of the United States of America 114 (16). National Academy of Sciences: 4165–70. doi:10.1073/pnas.1613616114.

Sah, Pratha, Lisa O Singh, Aaron Clauset, and Shweta Bansal. 2014. “Exploring Community Structure in Biological Networks with Random Graphs.” BMC Bioinformatics 15 (220). BioMed Central. doi:10.1186/1471-2105-15-220.

Salathé, Marcel, and James H Jones. 2010. “Dynamics and Control of Diseases in Networks with Community Structure.” Edited by Christophe Fraser. PLoS Computational Biology 6 (4). Public Library of Science: e1000736. doi:10.1371/journal.pcbi.1000736.

#### **CHANGE-LOG:**

- First stage of project found little variation in M:F case ratio with assortativity only so next stage will compare the effects of sex-specific susceptibility and assortativity
- SIR model on networks died out fairly quickly, not representative of TB in populations so next stage will compare SIR and SIRS models on networks
- First stage of project found little variation in disease spread on networks of different sizes so next stage will focus on networks with 1000 nodes