PROTOCOL FOR:

Sex-assortative contact patterns and male-bias of Tuberculosis

Last updated: Sep 10, 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 differences in how men spread 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 but there is a lack of evidence for this (Mossong 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 preferential mixing by sex (i.e., sex-assortativity). Other hypotheses exist for explaining male-bias including higher male susceptibility due to immunological differences (Nhamoyebonde et al. 2014), higher male transmissability due to larger cough volume (Lindsley et al. 2012, Gupta et al. 2009), and longer male infectious period due to delaying care-seeking (Borgdorff et al. 2000). Whether and how aspects of social mixing patterns and biology interact to drive male-bias in TB has yet to be determined.

The effects of assortativity (i.e., modularity) on epidemic dynamics (e.g., total outbreak size or equilibrium prevalence) have been studied with math models. Results indicate that the effect of assortativity depends on assumptions about recovery and supply of susceptible individuals (SIR, SIS) as well as the distribution of subgroup size. 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 (Nadini 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 (Nandini 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 vary in their capacitity to spread infection.

In this project, we will investigate whether sex-assortativity in social networks can contribute to male-bias or sex-specific differences in male disease transmission are necessary to explain male-bias. In addition, we will compare how assortativity changes epidemic trajectories in different disease systems. SIRS models have the ability to capture the persistance of TB in endemic areas.

Research questions:

1. Can sex-assortative mixing lead to male-bias by itself or are alternative male-bias explanations (susceptibility, transmissibility, infectious period) required to explain male-bias?
2. What are the effects of sex-assortative mixing and alternative explanations (susceptibility, transmissibility, infectious period) on disease spread (peak size/time, variation in outbreak size/duration)?

Study design:

We will examine the effects of sex-assortativity and alternative explanations on the ratio of male to female cases using SIR and SIRS models of disease spread on synthetic contact networks.

Alternative explanations are:

1. Higher male susceptibility modeled as higher probability of infection if the target node is a male
2. Higher male transmission modeled as higher probability of infection if the source node is a male
3. Longer male infecitous period modeled as longer average duration of infection

Network generation:

In the synthetic contact networks, each node will represent an individual and each link is an interaction between individuals that infection can spread.

To understand how increasing assortativity could change patterns of sex-bias in TB, we will will vary assortativity, r. Assortativity will be 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 $-1 \ll r \geq 1$ 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 ϵ of r, randomly choose a proportion α 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 α 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. Parameter range will be extended for publication.

Variable	Value
Sex-assortativity, r	(0, 0.9) by 0.3
Degree distribution, $p(K)$	$\frac{k^{-\alpha}}{\zeta(\alpha)}$
Mean degree, $< K >$	10
Network size, N	$1 \cdot 10^3$
Tolerance, ϵ	0.035
Rewiring proportion, α	0.2
Replicates	300

Models of disease transmission:

An infection spreads along an edge at probability depending on the overall transmission rate, β. Overall transmission is determined by baseline infection rate τ, the transmissibility of the source node (α_t), the susceptibility of the target node (α_s), and the duration of infection, (T) so that $\beta \sim \tau \cdot \alpha_t \cdot \alpha_s \cdot T$. T, the realized infectious period for a node, depends on α_i. The probability that x transmits to y is $1 - e^{\beta}$. Infected individuals recover at an exponentially distributed recovery rate γ which can vary between male and female nodes (see below).

Other transitions are independent of node sex. Once infected, latent individuals become infectious at an exponentially distributed recovery rate δ. Recovered individuals revert to susceptible at at an exponentially distributed recovery rate ψ. Parameter values for models are given in Table 2.

Table 2: Disease parameters for models. Parameter range will be extended for publication.

Variable	Value
Initial susceptible, S_0	$N - 0.05 \cdot N$
Initial latent, L_0	0
Initial infected, I_0	$0.05 \cdot N$
Initial recovered, R_0	0
Baseline infection rate, τ	0.025, 0.045, 0.08, 0.12
Latent period, $1/\delta$	4
Infectious period, $1/\gamma$	2
Reversion rate, ψ	0, 0.1
M:F susceptibility ratio, α_s	1, 1.5, 2
M:F transmissibility ratio, α_t	1, 1.5, 2
M:F infectious period ratio, α_i	1, 1.5, 2

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) and will be implemented with an event-based algorithm. Male:female differences in transmission α will alter the overall transmission rate depending on the source and target node as follows:

$$x_f = \frac{2}{\alpha + 1}$$

$$x_m = \frac{2\alpha}{\alpha + 1}$$

where x is either τ, or ω,and f and m represent female and male nodes and α represents the parameter ratio between males and females. These equations satisfy

$$0.5x_m + 0.5x_f = 1$$
$$x_m = \alpha x_f$$

so that the average x is 1.

Table 3: Example transmission rates by sex for male susceptibility and transmissibility scenarios. Given these formulas, when men are the targets, rates are higher. The same rate changes will be performed for male transmissibility except when men are the "sources", the rate will be higher.

Source -> Target	Overall transmission rate, β
F->F	$\frac{\tau \cdot 2}{(\alpha+1)}$
M->M	$\frac{\tau \cdot 2 \cdot \alpha}{(\alpha+1)}$
M->F	$\frac{\tau \cdot 2}{(\alpha+1)}$
F->M	$\frac{\tau \cdot 2 \cdot \alpha}{(\alpha+1)}$

Longer male infectious periods will be incorporated in largely the same way, but solving for $\frac{1}{\gamma}$ is slightly different. Instead of solving the above equations, I will use

$$x = \frac{c(d+1)}{2}$$
$$y = \frac{c(d+1)}{2 \cdot d}$$

where $x = \gamma_f, y = \gamma_m, c = \gamma, d = \alpha$.

Kiss, Miller, & Simon (2017) suggest α_s is independent of R_0 while α_t is not. They don't discuss the relationship between R_0 and α_i.

Parameters for τ were identified with the following analytical expression for R_0 in a continuous-time Markov model for SIR diseases (see Kiss, Miller, and Simon, page 221):

$$\frac{\tau}{\tau + \gamma} \frac{< K^2 - K >}{< K >} > 1$$

Numerical validation for the solution will be obtained by simulating model with transmission rates leading above and below $R_0 = 1$.

Analysis:

Network structure:

- Validation of network structure will include checking mean degree, number of edges, number of nodes, connectivity, presence of multiple edges, self-loops, etc.

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

M:F ratio:

- Effect of assortativity on M:F case ratio
- Effect of male-susceptibility on M:F case ratio
- Effect of male-transmissibility on M:F case ratio
- Effect of male-infectious eriod on M:F case ratio
- Interaction between assortativity and alternate explanations on M:F case ratio
- Variation of results by model type (SIS, SIR, SLIR, SIRS, SLIRS)
- Variation of results by transmission rate

Disease spread:

- Effect of assortativity and alternative explanations on probability of an outbreak, outbreak size, duration
- Variation of results by model type and transmission rate
- Prevalence of LTBI across simulations

Model fit:

- TBD: Prevalence of LTBI and male-bias across parameter combinations

Checklist:

- X Figure out how to code differing infectious periods
- X Figure out overall transmission rates when transmissibility AND susceptibility vary
- X Modify model script to handle latent class
- X Modify model script to handle male transmissibility and infectious period
- X Write additional analysis code to test for the effects of alternative hypotheses on male-bias
- X Interpret new results
- Repeat analysis with small-world networks?
- Focus on parameterization of the models since parameterization is arbitrary at the moment, focusing on fitting prevalence of latent infection in a population (25-30%)
- Analyze the dual effects of individual heterogeneity on infection?

Important background papers:

Craft, Meggan E. 2015. "Infectious Disease Transmission and Contact Networks in Wildlife and Livestock." Philosophical Transactions of the Royal Society B: Biological Sciences 370 (1669). The Royal Society: –20140107. doi:10.1098/rstb.2014.0107.

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

Lindsley, William G, Terri A Pearce, Judith B Hudnall, Kristina A Davis, Stephen M Davis, Melanie A Fisher, Rashida Khakoo, et al. 2012. "Quantity and Size Distribution of Cough-Generated Aerosol Particles Produced by Influenza Patients During and After Illness." Journal of Occupational and Environmental Hygiene 9 (7). Taylor & Francis Group: 443–49. doi:10.1080/15459624.2012.684582. **This paper measures the volume of coughs produced by 23 subjects with influenza and supplementary data suggest that men have higher cough volume than women (though the paper wasn't focusing on gender). Further analyses also show correlation cough volume with weight and height. An earlier paper by the same first author also published cough volume data for influenza patients and found men to have nearly twice the volume of women on average.**

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:

- 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 (July 6)
- Found little variation in disease spread on networks of different sizes so next stage will focus on networks with 1000 nodes (July 6)
- Found results to be sensitive to transmission rates, next stages will focus on epidemics with varying transmission rates (July 24)
- Following committee meeting, no one really pushed for including births and deaths in network. Instead, wanted to include additional hypotheses for male-bias in analysis. Edited protocol to include new parameterization and checklist (Sep 9)
- Changed protocol to assess simulated prevalence of LTBI in addition to male-bias to identify "best models".