Use of Network Analysis to Model the Effect of HIV PrEP on the Spread of HIV and Gonorrhea

Giovanni S. P. Malloy and Rodrigo Silva-Lopez

Abstract

To fully comprehend how HIV spreads over a community, and what kind of measures can be implemented to control its spreading, it is important to model all the different connections existing between its members. Many existing models of disease spread fail to do so. One way to model these connections is through the use of network analysis. Although there are several studies that have dealt with this task, they have lacked two main features: First, the intervention to prevent disease spread is not usually compatible with the results of an optimization performed using network techniques, and secondly, previous studies have neglected the interaction between HIV and other sexual transmitted diseases (STDs). In this report, we find that HIV Pre-Exposure Prophylaxis (PrEP) may reduce the prevalence of gonorrhea in MSM communities. Additionally, we look at several different edge deletion methods to prevent the spread of HIV. Among these, the minimization of connectivity between HIV susceptible and HIV infected individuals is most promising for preventing disease. We aim to model this question for a network of men who have sex with men (MSM) in an urban setting in the United States.

1. Introduction

HIV has been endemic in the United States for multiple decades. In 2012, the United States Food and Drug Administration first approved pre-exposure prophylaxis, PrEP, for the prevention of HIV [16]. PrEP is a pill that individuals take once per day to reduce the risk of HIV infection [16]. This drug was originally heralded as the beginning of the end of HIV and AIDS. However, some studies suggest that MSM individuals who take PrEP are more likely to engage in risky sexual behavior and contract other sexually transmitted-infections, such as gonorrhea [4, 5]. We aim to model the relationship between HIV and gonorrhea in MSM communities using social network analysis. Given the importance of network structure to disease spreading, it is natural to consider whether and how network information can be used to design more effective disease control intervention. One way to achieve this is the use of algorithms for link removal in the network as a way to control the spreading of diseases. In general, as observed in the paper presented [2], the disease network can be modeled by an undirected adjacency matrix with specific characteristics for the disease. Given the nature of the phenomena, this network must evolve as a function of time. For the outbreak of the disease, a susceptible – infected – recovered (SIR) model is used, where a node become infected at a rate $\beta$ and nodes recover at a rate $\delta$. In this case, the
algorithms are centered in providing the best option to remove $K$ edges.

2. Literature Review

We will explore three papers as a literature review for our work. The first paper lays the foundation of a generic framework for network modeling for HIV and other sexually transmitted diseases. However, this paper does not consider the use of PrEP or the interacting effects of sexually-transmitted diseases. The second two explore different algorithms for edge deletion in a network with disease spread. Edge deletion is an important concept in disease spread, as it serves as the analog to policy interventions to mitigate disease spread.

2.1. Enns, Brandeau (2011)

Often, infectious diseases are modeled with assumptions of homogeneous mixing in the population. However, sexually-transmitted diseases benefit greatly from modeling specific connections [1]. By including a greater degree of granularity in modeling approach, such an analysis would be more useful to health policy decision makers. Enns and Brandeau [1] lament the difficulty in finding specific sexual contact networks, so they propose a generalized framework for sexual contact network generation. They use a similar algorithm to the Erdős-Rényi random graph. However, instead of a network with a Poisson degree distribution, the degree distribution is determined by sexual contact surveys where participants are asked about the number of sexual partners they have over a period of time. In this way, [1] can create a representative model of a sexually active population without having access to explicit data. Not every type of sexual contact is the same. This is made explicit in the adjacency matrix, where spousal partners are given weight 1 and non-spousal partners weight 2.

2.2. Enns, Brandeau (2015)

Enns and Brandeau [2], proposed four different link removal approaches which can be classified according to their occurrence: Before (Preparation) or after the event (Reaction), or by their strategy, which can be either rank-based or optimization-based. The four methodologies are all the possible combinations of these types, being: Removing links in order of edge centrality (prevented, rank-based), removing links in order of susceptible - infected edge centrality (reactive, rank-based, $R_0$ minimization (preventive, optimization-based), and optimal quarantining (reactive, optimization – based). Given that the HIV and gonorrhea diseases are already affecting the population, we will only focus on reaction methods. Regarding Rank-based methods, for the reaction phase, the rank of an edge is:

$$c_{st}(e) = \sum_{(i,j) \in E} \sum_{s_0(i) = 0, s_0(j) = 0} \frac{\sigma(i,j,e)}{\sigma(i,j)}$$

Where $\frac{\sigma(i,j,e)}{\sigma(i,j)}$ is the fraction of the shortest paths between nodes $i$ and $j$ passing over a link $e$, but the sum is only performed over the shortest paths between infected and susceptible nodes. In this case, given a limited amount of resources, $K$ edges will be eliminated from the network. Note that this approach assumes that the budget is not big enough to delete the trivial solution of the worst $K$ nodes.

2.3. Nandi, Medal (2015)

Another paper that uses link deletion as a measure to control diseases in the one developed by Nandi and Medal [3]. These authors proposed four different methods that minimize the spread of a disease in terms of the elimination of edges, which is different than monitoring the spread of the disease as links are eliminated. These four different methods have different objective functions, which are defined
by: 1- Minimizing the number of pairwise connections between infected and susceptible nodes. 2- Minimizing the number of susceptible nodes at risk of infection. 3- Maximizing the number of transmission paths removed from the network and 4- Minimizing the total weight of transmission paths between all of the infected nodes and all of the susceptible nodes. Since the proposed models can be complicated or computationally expensive, the authors proposed heuristics version of each algorithm.

3. Motivation

Modeling the HIV disease is not enough. We now know that there may be negative effects on the control of diseases like gonorrhea with increased use of PrEP [4, 5]. These diseases ought to be accounted for when describing the true public health impact of PrEP distribution. Instead of rigorous network analysis techniques as a computational proxy for a public health intervention, [1] uses an approximate empirical result.

4. Methods

4.1. Data

There are three main data considerations for this project. The first is a distribution of sexual contacts stratified by both casual and steady sexual partners. This can be found in literature detailing the results of sexual contact surveys among MSM populations (Table 1). This information was taken directly from literature. While some of the data are probabilities, some are rates per person-year, and some are rates per person-0.5 year, they are all transformable to the appropriate form for the model.

The second data stream is information on partner formation and dissolution, which is approximated using average partner duration. There are typically two types of partners: ca-
Table 2. Partnership Information

<table>
<thead>
<tr>
<th>Duration of Partnership</th>
<th>Percentage</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Casual Partners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 month</td>
<td>55.80%</td>
<td>[15]</td>
</tr>
<tr>
<td>1-6 months</td>
<td>21.60%</td>
<td>[15]</td>
</tr>
<tr>
<td>7-12 months</td>
<td>7.30%</td>
<td>[15]</td>
</tr>
<tr>
<td>13-24 months</td>
<td>7.40%</td>
<td>[15]</td>
</tr>
<tr>
<td>25-36 months</td>
<td>2.30%</td>
<td>[15]</td>
</tr>
<tr>
<td>&gt;37 months</td>
<td>5.70%</td>
<td>[15]</td>
</tr>
<tr>
<td>Steady Partners</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;1 month</td>
<td>17.70%</td>
<td>[15]</td>
</tr>
<tr>
<td>1-6 months</td>
<td>29.00%</td>
<td>[15]</td>
</tr>
<tr>
<td>7-12 months</td>
<td>14.40%</td>
<td>[15]</td>
</tr>
<tr>
<td>13-24 months</td>
<td>15.00%</td>
<td>[15]</td>
</tr>
<tr>
<td>25-36 months</td>
<td>8.30%</td>
<td>[15]</td>
</tr>
<tr>
<td>&gt;37 months</td>
<td>15.70%</td>
<td>[15]</td>
</tr>
</tbody>
</table>

Table 3. Initial Disease State Conditions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate of HIV I (no PrEP) to HIV I (No Tx)</td>
<td>.0086 per person-year</td>
<td>[10], Calculated</td>
</tr>
<tr>
<td>Rate of HIV S (PrEP) to HIV I (No Tx)</td>
<td>.0048 per person-year</td>
<td>[10], Calculated</td>
</tr>
<tr>
<td>Rate of HIV I (No Tx) to HIV I (Tx)</td>
<td>.43 per person-year</td>
<td>[11], Calculated</td>
</tr>
<tr>
<td>Rate of Gonorrhea S to I given HIV S (no PrEP)</td>
<td>0.211 per person-year</td>
<td>[5]</td>
</tr>
<tr>
<td>Rate of Gonorrhea S to I given HIV S (PrEP)</td>
<td>0.2844 per person-year</td>
<td>[5]</td>
</tr>
<tr>
<td>Rate of Gonorrhea S to I given HIV I (no Tx)</td>
<td>0.211 per person-year</td>
<td>Assumed</td>
</tr>
<tr>
<td>Rate of Gonorrhea S to I given HIV I (Tx)</td>
<td>0.2558 per person-year</td>
<td>[5]</td>
</tr>
<tr>
<td>Initial Proportion HIV I</td>
<td>0.108</td>
<td>[10]</td>
</tr>
<tr>
<td>Initial Proportion HIV S (PrEP) given S</td>
<td>0.395</td>
<td>[6]</td>
</tr>
<tr>
<td>Initial Proportion HIV I (Tx) given I</td>
<td>0.76</td>
<td>[12]</td>
</tr>
<tr>
<td>Initial Proportion Gonorrhea I</td>
<td>0.06</td>
<td>[13]</td>
</tr>
</tbody>
</table>

sual and steady. As shown in Table 2, the literature does not define these relationships by duration of partnership but rather by arbitrary distinctions drawn by survey respondents.

Lastly, information regarding natural history of disease models will be necessary. Some sexual contact surveys include prevalence of disease among the survey participants’ partners (Table 3). In order to focus our efforts primarily on the network analysis component of the project rather than the disease model, we will use a relatively simple model similar to that proposed by [8]. In this model, the population is stratified by gonorrhea and HIV disease state. We will consider both susceptible and infected states for gonorrhea, as re-infections are possible. For HIV, we will model using susceptible, infected, infected and on treatment, AIDS, and AIDS and on treatment. For all states, we will include a simple measure of PrEP status based on acceptability and motivations for adherence [9]. In order to calculate the rate at which HIV susceptible individuals (for those both on PrEP and not on PrEP) become infected, we used the overall HIV incidence rate for the MSM population in the US [10], the proportion of MSM individuals willing to take PrEP [6], and the reduction in HIV incidence rate due to PrEP [6]. To calculate the rate of seeking treatment for HIV, we used a probability of HIV re-testing for MSM populations and assumed that if a person was infected and got tested, then they would start treatment [11]. Finally, we assumed that the incidence rate of gonorrhea for someone who is HIV infected but not on treatment is the same as that of someone who is HIV susceptible and not on PrEP because they would not be aware of their HIV status.
4.2. Network Generation

As a fundamental part of this work, a network that captures both the effects of spreading of gonorrhea and HIV had to be defined. The network was created using a graph configuration model. We generate a series of \( n \) node degrees \( k_i \) based on the distribution of sexual partners found in Table 1. For each node \( i \) we create a box of \( k_i \) sub-nodes and randomly connected all of these sub-nodes. We then collapse each box into a supernode. As \( n \) gets sufficiently large, we expect that the number of self-edges and multi-graph edges will approach zero. This was be performed in a similar way to the work developed by [1], in which they probability of different nodes being connected is given by a specific probability distribution. The source of this distribution comes from information on sexual contact surveys [5, 6, 7]. In terms of the size of the network, initially our network uses one 1000 nodes, but its size is subject to the computational cost of implementing the edge deletion algorithms. For this project, only men that have sex with men will be considered.

Considering that the network represents the spread of the disease, it cannot be modeled as a static network, but it changes in time. At each time step, four basic processes will take place. First, relationships will form or dissolve using the data from Table 1 and Table 2 along with the methods proposed in [1]. The node distribution represents the distribution of steady and casual partners represented by the sexual contact survey data. Dissolution of relationships occurs as a function of the inverse of partnership duration. As each edge is created, it assigned an attribute of either casual or steady and given this attribute, assigned a partnership duration. Second, policy interventions was implemented in the form of edges deletion algorithms. Third, HIV infected individuals potentially progress to treatment based on the re-testing rates (Table 3). Finally, the sexual interactions occur. Only risky sexual interactions lead to potential disease transmission. For this project, we consider both transmission of HIV and gonorrhea, however, we do not discriminate between oral, anal, or urethral gonorrhea. The probability of spreading a disease from a node depend on the disease state of an individual regarding the diseases. In particular, each node has an attribute regarding HIV disease state (susceptible on PrEP, susceptible not on PrEP, infected on treatment (Tx), and infected not on treatment) and a gonorrhea disease state (susceptible or infected).

The initial condition of having HIV or Gonorrhea was considered independent, which means that there is no initial correlation between having HIV and Gonorrhea. The number of people infected with each disease is available in Table 3. This underlying prevalence has the potential to heavily influence the spread of the disease and is inherently uncertain. Different levels of PrEP prevalence will be used to measure its effect in the relationship between gonorrhea and HIV. However, we use a standard proportion of 0.395 in the base case [6].

4.3. Link Deletion Implementation

We used two different approaches to implement edge deletion algorithms. First, we removed links in order of susceptible - infected edge centrality. This method was used successfully by [2] to isolate susceptible individuals from infected individuals (Figure 1). After implementing the exact algorithm, we were restricted by the computational expense. Therefore, we used an edge betweenness centrality ranking to sort the edges in the network and then only considered those edges which were susceptible-infected links. This method is akin to a reactive policy based on an individual’s response to a sexual contact survey.
Second, we considered an algorithm which eliminates the edges which have the maximum number of paths between susceptible and infected nodes (Figure 2). This was also computationally expensive, so we utilized a built-in function for edge load centrality in NetworkX. The load centrality of an edge is the fraction of all shortest paths that pass through that edge. Finally, we analyzed a random deletion method as a status quo comparison for the other edge deletion methods.

Since these models capture the interaction between HIV and gonorrhea, the edge deletion methods were implemented by using different weights on each disease to measure the importance of the edge when minimizing the total weight of transmission, or to perform random edge deletions. Given that there is no existing information on the values of this weights, different values were used. The link deletion process was at each step of the spread of the disease, and since there is no information on $K$ (number of edges to delete per step), a set of different $K$ was used.

4.4. **Evaluation**

As a measure of the performance of each edge deletion method, the authors use the number of patients infected at the end of the $N$ step. The value of $N$ for this analysis was ten.

5. **Results**

As a result of this project, this team will show the comparison between different edge deletion methodologies in terms of the number of people infected and computational costs. These results will be deaggregated by the type of disease. In addition, in order to show the efficiency of the methods, a comparison with random edge deletion and with the benchmark of no action, will be performed. The previous analysis will be performed for different values of $K$ (Edge deletion number). Finally, given the use of different PrEP use, the relationship between PrEP, and prevalence of gonorrhea and HIV will be shown. In terms of this milestone, besides setting up the network with its initial conditions, and its dynamics, this team developed a measure of how the name of infected people changed with the initial proportion of people on PrEP and some preliminary
deletion edges methods. Since the analysis is stochastic, several simulations were averaged on each experiment to get stable results.

5.1. Network Statistics

The network we created using our generation method matches our data very well. According to the data collected (Table 1), the average number of partners of an MSM individual over six months is 4 [14]. The network generated for this project had an average node degree of 4.5. This includes a mean of 1 steady partner and 3.5 casual partners every six months. One of the major assumptions of this model is a static degree distribution over the entire population (Figure 3). That is to say, the degree of individual nodes may change over time, as seen in Figure 4 fell and then rose again. It is important that these demographic features of the network match the data we obtained from the sexual contact surveys. Finally, the model predicts that over time, the number of individuals with only HIV or only gonorrhea drops while the number of individuals with both diseases increases over time (Figure 5).

5.2. Effect of PrEP

PrEP reduces the number of cases of HIV in the network and also decreases the number of individuals with gonorrhea (Figure 6). This is not the result that we expected to find given
the initial hypothesis of the effects of PrEP on the prevalence of both diseases in the MSM community.

Additionally, we look at the impact of casual partners and steady partners in the transmission of both HIV and gonorrhea. In general, it is difficult to decouple the influence of the type of relationship because steady relationships last longer, but casual relationships are more common. Over the course of our simulation, casual partnerships resulted in more cases of HIV than steady partnerships (Figure 7). In general we can see that casual partnerships result in more cases of HIV. This is expected because individuals have many more casual partners than steady partners.

5.3. Effect of Edge Deletion Methods

As mentioned before, initially two deletion algorithms were implemented. The first method deleted edges randomly, the second deletion method deleted edges by HIV susceptible-infected degree centrality, and the third minimizes connectivity of susceptible and infected individuals (Figure 8). The edge deletion policy that prevented the most cases of HIV was connectivity minimization. This algorithm was a significant improvement over that of random edge deletion.

5.4. Sensitivity Analyses

Many of the model parameters were uncertain. To determine the influence of these parameters on the results of the model, we ran several one-way sensitivity analyses. The parameters that we hypothesized to have the most influence on the outcome of the model were the efficacy of PrEP in preventing HIV (Figure 9) and the proportion of steady versus casual partners (Figure 10).

These parameter values represent a reasonable and plausible range. The probability of transmission of HIV for someone on PrEP can range from no effect to completely effective. Likewise, individuals are not likely to have
many steady partners, but the number of casual partners they may have at the same time as a steady partner will vary. Based on these analyses, we can see that PrEP efficacy and proportion of steady partners have a large influence on disease incidence of HIV. A higher efficacy of PrEP and higher proportion of steady partners leads to the lowest prevalence of HIV.

6. Discussion

The results of this project highlight the need for further investigation on the effects of HIV pre-exposure prophylaxis on the spread of sexually transmitted infections, such as gonorrhea. First, the two disease are at least slightly correlated. Over the course of the model, the number of individuals with both HIV and gonorrhea increased while the number of individuals with only HIV or only gonorrhea decreased. The hypothesis of the public health community is that these diseases are linked due to the increased prevalence of PrEP leading to riskier sexual behavior among MSM individuals. However, the results from this network analysis model show that the opposite might be true. PrEP appears to reduce the number of MSM individuals with HIV. This is a strong assertion to make given the prevailing knowledge of HIV and STI transmission. Therefore, further exploration of this relationship is required before coming to concrete conclusions. One early hypothesis of ours is that perhaps MSM individuals who take PrEP are more conscious of preventing sexually-transmitted diseases in general.

Another important result of this model is the importance of the type of partnership that dominates and individual’s life: steady or casual. Steady partners appear to be a safer option for MSM individuals looking to prevent the spread of disease. Given that individuals tend to have more casual partnerships, the spread of HIV and gonorrhea continues to
grow throughout the model despite the presence of PrEP. Both partnership type and PrEP efficacy had major effects on the results of the model.

The edge deletion method that showed the most promise was the minimization of connectivity of susceptible and infected individuals. Without a full network structure in mind, it would be difficult for health policy makers to find a real-world policy that mirrors this network analysis implementation.

A major limitation of our model is its reliance on self-reported sexual contact data. This is not always accurate and any theoretical implementation of a network model based on this limited amount of information could misrepresent the true underlying MSM sexual network. Second, the data collected for this project assumes a largely homogeneous population within a particular disease state. While node degree and type of partnership per edge will vary within these sub-populations, probabilities remain deterministic across the sub-population of a particular disease state. Finally, this project considers only the public health impact of the number of cases of HIV and gonorrhea and does not consider costs.

7. Conclusions

Based on the results of our MSM sexual network model, we conclude that PrEP has decreases the spread of gonorrhea while preventing the spread of HIV. This is an important departure from what is commonly hypothesized. When it comes to decreasing the prevalence of HIV, from a network analysis perspective, the most effective policy intervention to prevent the spread of HIV is connectivity minimization. Therefore, health policy makers should seek to implement a policy intervention that is most analogous to this type of edge deletion. By doing so, there is the potential to prevent cases of HIV and cases of gonorrhea.

The model results were most sensitive to the efficacy of PrEP and the proportion of steady partners to casual partners. This means that the value of information on both of these parameters for health policy makers is high. Our network model of an MSM community in the United States could be an important tool to inform good health policy decisions and more work should be done to improve it.

8. Future Work

Moving forward, we plan to address many of these simplifying assumptions by collecting better data and collaborating with experts in HIV at Stanford, such as Prof. Margaret Brandeau and Prof. Eran Bendavid. First, we aim to expand the size of our network model. The computational cost and time pressures of a quarter long project led to a small social network. Second, we hope to add cost assessments to ascertaining the information required to implement each edge deletion policy. With the new cost figures and the effectiveness of each edge deletion policy from our analysis, we can conduct a cost-effectiveness analysis. This is a common tool used by health policy experts in order to make decisions. Overall, the model presented in this project is a useful tool for determining the relationship between PrEP and HIV and gonorrhea incidence in MSM communities in the United States.

9. Code

https://github.com/malloyg32/cs224w

References


