Adamawa State University Journal of Scientific Research
Volume 10 Issue 1, 2022; Article no. ADSUJSR 1001011
ISSN: 2705-1900 (Online); ISSN: 2251-0702 (Print)
http://www.adsujsr.com

Modeling a Sexual Network using Graph Theory and the Simulation of Prevalence and Spread of Human Immunodeficiency Virus (HIV) in the Network

Kamal Bakari Jillahi[1]*, Elkanah Obadiah Sambo[2] and Aliyu Umar[3]
[1,3]Department of Mathematical, Sciences Taraba State University, Jalingo.
[2]Department of Biological Sciences, Taraba State University, Jalingo.
Contact: kamalbakari@gmail.com.

(Received in February 2022; Accepted in July 2022)

Abstract

The application of mathematical methods in epidemiology cannot be over emphasized. An area which has seen a remarkable development recently is the modeling, simulation and analysis of contagion networks. On the other hand, HIV/AIDS have remained endemic especially in sub-Saharan Africa spreading primarily through sex. This calls for a critical study of the contagion network of the most sexually active portion of the population to understand the structure and behavior of this network. This research modeled the sexual networks of some selected sex workers and their clients using the mathematical graph theory. Furthermore, using statistics obtained from two HIV clinics and the National Action Committee on Aids, the research simulated the prevalence and spread of HIV/AIDS in the modeled networks. Results indicate that in agreement with the widely held notion of scale-free, the networks have long-tailed power-law degree distribution with an exponent $\gamma = 2.9$ *(0.3) for females, and* $\gamma = 2.7$ *(0.2) for males and tend to be concentrated around sex workers with younger age who do not live in brothels. Furthermore, the observed networks exhibit a low variance in the number of contacts most sex-workers have over the given time period the study was carried out.*

Keywords: Simulation; Modeling; Contagion Network; Graph theory; Scale-free; Small-world

Introduction

A population, like any other complex entity, can be modeled as a network, in mathematical terms a graph; whose vertices are the single individuals and the edges represent the interactions among those vertices (Meyers *et al*., 2005). Since infectious diseases spread by interaction between infectious and susceptible people, the structure and behavior of the network of these nodes and edges plays a crucial role in the dynamics of the infectious disease propagation (Jones and Handcock, 2003; Bauch and Rand, 2000; Diekmann and Heesterbeek, 2000).

The overwhelming majority of people living with Human Immunodeficiency Virus (HIV) are on the Sub-Saharan Sub-Continent (Latora *et al.*, 2006) due to, (1) little or no understanding of the effects of some sexual behaviors by the people, (2) less efficient efforts in the fight against the spread of this disease (Latora *et al.*, 2006) as a result of lack of understanding the dynamics and behaviors of the social networks through which this disease spreads (Ferguson *et al.*, 2001),(3) Most studies on the dynamics of epidemics are population based which adopt over simplifying assumptions that make their applicability less acceptable (Latora *et al*, 2006; Jones and Handcock, 2003; Liljeros *et al.*, 2001; Ghani *et al.*,1998; Morris and Kretzschmar, 1997).

To this end, this work provides an insight into the dynamics of the sexual network of the subjects and studies the social, economic and biological factors responsible for the evolution and decay of these networks. Furthermore, we simulate the prevalence and spread of HIV in the modeled network using both primary data and statistics obtained from secondary sources in other to predict the prevalence and spread of the disease in the future under different population demographics, protection, therapy and vaccination regiments.

It is our believe that understanding and monitoring the dynamics of the contagion network over time will help in providing proactive solutions which may include suggesting the best quarantine regiment, isolation mechanism, targeted campaign against the spread of the disease.

In a network, the network density, the presence of sub-structures and the existence of long range connections typical of a class of networks named small worlds (Ghani and Garnett1998; Keeling, 1999; Liljeros *et al.*, 2001; Latora *et al*., 2006), increases the probability of secondary infections, enhancing critically the spread of infectious diseases. Moreover, the key role of central individuals (Meyers *et al*., 2005) has long been known, and epidemiologists have taken it into account in the modeling of disease spreading (Ferguson and Garnett, 2000). In most studies, (Ghani and Garnett1998; Keeling, 1999; Liljeros *et al.*, 2001; Latora *et al*., 2006;Ghani*et al*,1998; Morris and Kretzschmar, 1997) the emphasis is on the simplest measure of node centrality, the degree, and on the consequences deriving from the heterogeneity of the degree distribution of the network. The degree k of a node is the number of its neighbors, that is, the number of links adjacent to the node: in the case under study, the number of sexual partners of an individual. Recently it has been found that many biological, social and communication networks are different from random graphs (Hethcote and Van-Ark, 1987) and all share the same property of having a *long-tailed power-law degree distribution* $P(k) \cong k^{-\gamma}$ with an exponent γ ranging between 2 and 3;networks with such extremely heterogeneous degree distribution have been named scale-free networks. In such networks, due to the infinite variance of the degree distribution (Pastor-Satorras, and Vespignani, 2001), disease can spread and be maintained even when the infection probability (the transmissibility) is extremely small.

This has been proven initially for the spread of computer viruses over the Internet (Pastor-Satorras and Vespignani, 2001), although it can be relevant to understand the spread of pathogens in other types of networks, as networks of sexual contacts, that are important for the transmission of sexually transmitted diseases: herpes genitalis, gonorrhea, syphilis, Chlamydia, and HIV (Morris and Kretzschmar, 1997; Liljeros *et al*., 2003). It has been pointed out in literature that in such networks some individuals can have a large numbers of partners (Hethcote and Van-Ark, 1987; Keeling, 1999; Latora *et al.*, 2006). A recent survey in Sweden has shown that the network of sexual contacts is a scale-free networks and the numbers of different sexual partners over individual lifetime is a power law distribution with an exponent $\gamma = 3.1$ (0.2) for females, and $\gamma = 2.6$ (0.3) for males (Liljeros *et al*., 2001).

Materials and Methods

This research recruited ten (10) sex workers from three (3) different locations (three from Gindin-doruwa, four from Manga, three from Sabon-layi) all in Jalingo metropolis. These subjects agreed to provide information about their sexual contacts after every month for three months as a consideration for a payment at each data collection point. Then each of these sexual contacts was contacted either in person or via phone for their consent to participate in this study. Consenting respondents participate by filling out a questionnaire.

At each round of data collection, each of the sex workers and/or their client is issued a questionnaire which asked questions from personal contact (phone numbers), to sexual behavior, contacts (respondent may indicate a sexual contact or remain silent on nature of contact) in most situations where the respondents were unable to fill the questionnaire, the questions were interpreted to them and their response were entered in the questionnaire by a research assistant. After every round of data collection, the proposed model is thereby applied to the data and the simulation process performed. Where an individual with high centrality is identified or an abnormality in the structure of the network, an interview with the sex work(s) in the motif is arranged to verify received information.

The research used two different questionnaires to obtain data from respondents. The first questionnaire was designed to collect data from clients, this questionnaire asked questions such as phone number, age, current marital status, spousal information, number of spouses, work and residential address, partner preferences (physical appearances). While the second questionnaire was used to collect data from sex workers, this also asked similar questions but in addition it asked about casual partnerships and the rate at which individuals create or leave those partnerships.

Modeling the Sexual Partnership Network

In other to simulate the sexual network of the respondents, the collected data was represented as a 2 dimensional matrix $X_{i,\ j}$ where i=1,2,3,…N (total number of respondents) while j=1,2,3,…9 (attributes and responses of each respondent) where 1: mobile phone number, in situations where a respondent has more than one number, only one is considered and every other one is kept as an alias, 2: age in years, 3: gender $\in \{1,0\}$ where 1 represents a male and 0 represents a female, 4: place of work, 5: place of residence, 6: casual sexual behavior $\in \{1,0\}$ 1 where respondent engages in prostitution or patronizing prostitutes and 0 represent otherwise, 7: sex worker $\in \{1,0\}$ 1 where respondent is a prostitute and 0 otherwise, 8: condom usage $\in \{1,0\}$ 1 where respondent always uses condom and 0 otherwise, 9: HIV status $\in \{1,0\}$ this is not obtained from the respondent therefore initially set 0 for all nodes and updated as the prevalence and spread of HIV process continues (see section on Simulating the Prevalence and Spread of HIV in the Network). Furthermore, information about sexual contacts collected from the respondents is also represented as 2 dimensional matrix $A_{i,j}$ where i = 1,2,3,..N and j=1,2,3,…M represents the number of sexual partners of each respondent. The link weight between nodes of the network $EW_{i,j}$ is also represented as a 2 dimensional matrix where i = 1,2,3,..N represent the weights of relationship of the corresponding $A_{i,j}$ entry and j=1,2,3,…M entries of each paticipant.

Establishing a Link between Nodes

The sexual partnership between individuals is represented as a weighted graph. Assuming that all respondents are sexually active adults and the prevalence of homosexuality is negligible, the link weight EW between two individual nodes x_i and x_j is a number [0, 1] which is obtained using the system in (4):

$$EW(x_i, x_j) = \begin{cases} f_1, & x_i \in \{A_j\}\ AND\ x_j \in \{A_i\} \\ f_2, & x_i \in \{A_j\}\ XOR\ x_j \in \{A_i\} \\ f_3, & (x_i \in \{A_j\}\ XOR\ x_j \in \{A_i\})\ AND\ (x_i \notin \{X\}\ XOR\ x_j \notin \{X\}) \end{cases} \qquad (1)$$

Here

$$f_1 = x_{i,3} XOR x_{i,3}\ \times\ (w_1 \times\ x_i \in \{A_j\} + w_2 \times\ x_j \in \{A_i\}) \qquad (2)$$

where w_1, w_2 [0, 1] such that $w_1 + w_2 = 1$ which signifies the relevance of consent of node i and j respectively. From the system (1), the first part eliminates the possibility of homosexuality from the modeled network or ensures that only nodes with opposite gender can have sexual partnership. To further explain, this function is used where both nodes indicate having sexual partnership. Where either of the two nodes indicates partnership but the other does not reciprocate, f_2 is used to indicate partnership between the two nodes.

$$f_2 = x_{i,3} XOR x_{i,3}\ \times\ (w_1 \times\ x_i \in \{A_j\} + w_2 \times\ x_j \in \{A_i\}) +\ \omega_3 \times \beth + \rho/\Omega + \omega_4 x_{i,6} + \omega_5 \times \tau(x_{i,4}, x_{j,4}) + \omega_6 \times \psi(x_{i,5}, x_{j,5})) \qquad (3)$$

where $\beth(x_i, x_j) \in \{1,0\}$ signifies if x_i and x_j are in one another's contact, ρ is the number of common contacts between x_i and x_j, Ω is the total number of different contacts of x_i and x_j, $\tau(x_{i,4}, x_{j,4})$ and $\psi(x_{i,5}, x_{j,5})$ are measures of closeness between node i and j's place of work and residence respectively. τ and ψ can be computed using any distance measurement techniques such as Euclidian distance. In this work, each of the three study locations was represented using the scheme as follows;

(i) Gindin-mangoro,
(ii) Sabon-layi, and
(iii) Gindin-dorowa.

Therefore, a simple similarity measure was used for closeness such that if node i and j work or live in the same location they will have the same work or residence location. $w_1, w_2, w_3, ..w_6 [0, 1]$ Such that $\sum_{i=1}^{6} w_i = 1$ which signifies the relevance of either parties indication of partnership, common contacts, casual sexual behavior, place of work and place of residence respectively. Finally, where one node indicates sexual partnership with an individual who is not part of the respondents for the study the third function (f_3) is used to calculate the weight of the partnership.

$$f_3 = x_{i,3} XOR x_{i,3} \times \left(w_1 \times x_{i,7} + w_2 \times \frac{1}{x_{i,2}} \times w_3 \times x_{i,6} \right) \qquad (4)$$

where w_1, w_2, w_3 [0, 1] such that $w_1 + w_2 + w_3 = 1$ which signifies the relevance of casual sexual behavior, age and prostitution or patronage of node i respectively.

Simulating the Prevalence and Spread of HIV in the Network

In order to simulate the prevalence of HIV in the modeled network, Statistics from the National Action Committee on AIDS (NACA) website and two HIV clinics (Government house Clinic and FMC) all in Jalingo were used simulate the initial prevalence of the disease in the network. The prevalence of HIV among sex-workers Φsw was set at 48% while Φnf among non sex-workers females was 7%, the aggregate statistics of prevalence among general males (prostitute patronizing and non-patronizing) $\Phi sc = \Phi nm$ of 6% was used. The research initially designated a random number of nodes as infected based on their gender and sex-worker status. That is, 5 nodes among the sex-worker nodes, 2 nodes among the female non-sex worker nodes and 10 nodes among the male nodes. Assuming the spread of HIV through other means is negligible, usage of Human Anti-Retroviral Therapy (HART) does not confer lifelong immunity and recovery, the rate of contact between susceptible and infected is constant, the probability of transmission $PT(x_i , x_j)$ from an infected node x_i to an uninfected node x_j is set using the following function:

$$PT(x_i, x_j) = \begin{cases} p_1, & x_{i,8} = 1 \; AND \; x_{J,8} = 1 \\ p_2, & x_{i,8} \neq 1 \; XOR \; x_{J,8} \neq 1 \\ p_3, & Otherwise \end{cases} \qquad (5)$$

From the system in (5), p_1 is used in a situation where both the infected and uninfected nodes always use condom, while p_2 is used in the case where only one node uses condom and p_3 is used in the case where none of the nodes always uses the condom. To designate a previously susceptible node x_i as infected by an already infected node x_j, we use the conventional logistics activation function $y = \frac{1}{1+e^{-z}}$ where $z = PT_{i,j} \times EW_{i.j}.RANDNUM$. If $y>=$ some threshold then the HIV status ($x_{i,9}$) of the non-infected node is updated to 1.

Network/Infection Dynamics over time

As earlier noted, the structure of a network highly affects the possibility of spread of contagious diseases. Therefore to understand how the disease will behave in the modeled network, it is only logical to repeat the process over a period of time. In each time step the network is modeled using the collected data, the weight of links of nodes and probability of transmission is calculated using the system as in (1) and (5) above then finally the spread of HIV is simulated using the procedure as in section Network/Infection Dynamics Over. This can be algorithmically presented as below;

```
initialize X,A, threshold,$w_1$,$w_2$,$w_3$,$w_4$,$w_5$,$w_6$
EW=0
PT=0
designateΦsw nodes from sexworkers nodes as infected
designateΦnf nodes from non-sexworkersfemale nodes as infected
designateΦsc nodes from sexworkers clientmale nodes as infected
designateΦnm nodes from non-sexworkersclient male nodes as infected
$f_1(x_1,x_2)\{$
$return x_{1,3} XOR x_{2,3} \times (w_1 \times x_1 \in \{A_{x1}\} + w_2 \times x_2 \in \{A_{x2}\})\}$
$f_2(x_1,x_2)\{$
$return x_{1,3} XOR x_{2,3} \times (w_1 \times x_1 \in \{A_{x1}\} + w_2 \times x_2 \in \{A_{x2}\}) + \omega_3 \times \beth + \rho/\Omega + \omega_4 x_{1,6} + \omega_5 \times \tau(x_{1,4}, x_{2,4}) + \omega_6 \times \psi(x_{1,5}, x_{2,5}))\}$
$f_3(x_1,x_2)\{$
$return\ x_{i,3} XOR x_{i,3} \times \left(w_1 \times x_{i,7} + w_2 \times \frac{1}{x_{i,2}} \times w_3 \times x_{i,6}\right)\}$
fork=1 to timeframe
            i = 1
            foreach node do
                  get all partners of node i
                  j=1
                  foreach partner of node i do
                        if () then
                              $EW_{ij}$=$f_1$($x_{i,j}$,partner node)
                        else if () then
                              $EW_{ij}$=$f_2$($x_{i,j}$,partner node)
                        else if () then
                              $EW_{ij}$=$f_3$($x_{i,j}$,partner node)
                        end if
                        j++
                  endforeach
                  i++
            endforeach
            i=1
            for each infected node do
                        j=1
                  foreach partner of node
                        if ($x_{i,8}$=1 OR $x_{j,8}$=1) then
                              $PT_{ij}$= $p_1$
                        else if ($x_{i,8}$=1 XOR $x_{j,8}$=1) then
                              $PT_{ij}$= $p_2$
                        else
                              $PT_{ij}$= $p_3$
                        end if
                        z = $EW_{ij}$x $PT_{ij}$x Randnum
                        y = 1/(1+e^(-z))
                        if (y=>threshold) then
                              $x_{i,8}$=1
                        end if
                        j++
```

```
            endforeach
                i++
        endforeach
    update statistics {Φsw, Φnf, Φsc, Φnm}
    report Current result
    K++
end for
```

Result Discussion:

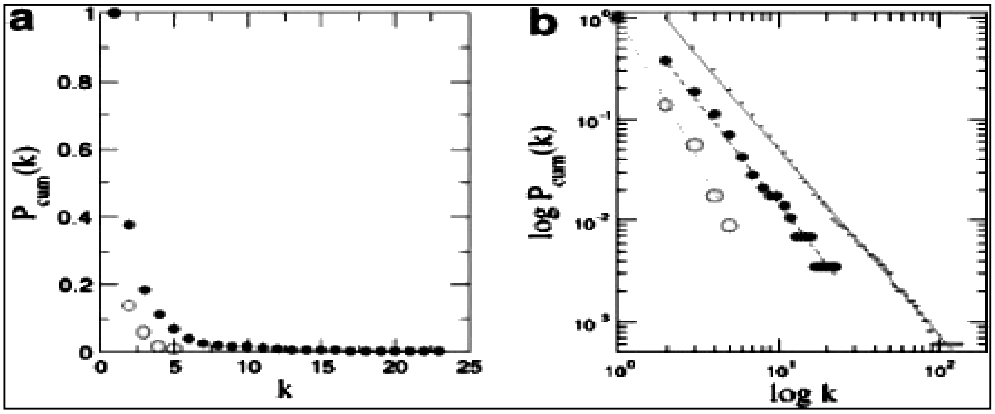

Figure 1: The cumulative distribution of the number of sexual partners in the previous 6 months is reported both for the33 females (both prostitutes and non-prostitute) (open circles) and for the 287 males (filled circles) in a linear scale (a) and in a log-log scale (b).

The observed network as modeled for the three months (see figure 2) is a scale free with $\gamma = 2.9$ (0.3) for females, and $\gamma = 2.7$ (0.2) with a monthly average of 5.94 ± 0.03 (Median 2, range 1 -18) for males and 28.5 ± 0.02 (Median 20, range 1 -36) for females. It should be noted that the population of female respondents is more diverse due to the presence of non-prostitutes (with average of 1.30 ± 0.001, median 2, range 1 -3) who are partners of the male clients of prostitutes. Figure 1(b) shows the cumulative degree for both male and female respondents with an artificial scale-free graph as comparison. Consequently, this implies that isolation mechanisms, quarantine regiments and prevention campaigns that target a few highly centralized individuals will be more successful than any other action like campaigns that targets social groups such as the diffusion of condoms among students at schools.

Furthermore, public health strategies for reduction and eradication of HIV based on reducing transmissibility, shortening the duration of infection and reducing the contact rate between susceptible and infected which aim at highly centralized individuals will be the best disruption strategy. In other words, programs as such frequent HIV test, economic empowerment and social orientation (which take them out of sex work) for sex workers with higher number of partners should be considered by stakeholders in the fight against HIV. In the same vein, since the usage of Human Anti-Retroviral Therapy (HART) reduces the transmissibility of infected individuals (Keeling, 1999) Campaigns to prioritize and encourage infected sex workers on the use of this drug should be highly encouraged among sex workers.

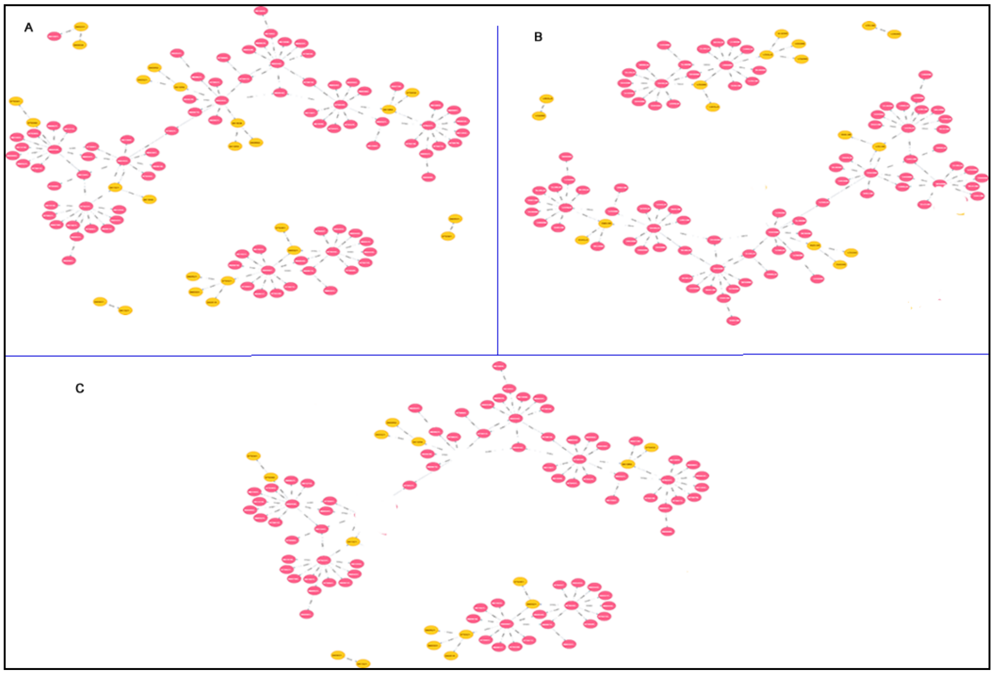

Figure 2: Sexual Network of respondents for the months of June (A), July (B) and August (C) 2018 showing un-married (orange) and married (red) nodes using the proposed simulation process

The cumulative degree distribution of the observed network from the collected data is asymmetric and extremely wide for both males and females (1-15 partners for males and 1-25 partners for females). Although the collected data cannot be used to draw conclusions on national and regional levels it is believed that the study area is replica of many locations regionally and nationally and thus conclusions can be applied across these areas with the same characteristics. This implies that males with 6 or more partners and females with 29 or more partners are prime targets for any campaigns that seek to reach highly connected individuals within these networks.

Using the Susceptible-Infected-Susceptible (SIS) epidemiological compartmental model which assumes recovery (un-transmissibility) in this case treatment using HART does not give life immunity (as infected individuals can either be susceptible or infected based on the usage of therapy) was used to calculate Susceptibility and Infectivity of individuals within the network using the systems in (6);

$$\frac{dS}{dt} = \frac{\beta SI}{N} + \gamma I \qquad (6)$$

Where β is the contact rate, which takes into account the probability of getting the disease in a contact between a susceptible and an infectious subject. S is the number of Susceptible, I is the number of Infected, N is the number in the population and γsimply the rate of recovery or death.

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I \qquad (7)$$

The variable parameters in the system (7) are the same as in (6). The change in susceptibility with respect to time $\frac{dS}{dt}$ was found to be ∞ which signifies that as time is kept constant the number of susceptible will be infinite. In the same vein the change in infectivity with respect to time $\frac{dI}{dt}$ was found to be ∞

Conclusion:

Knowledge of the structure of the sexual network provides important hints for public health organizations for control and eradication of sexually transmitted diseases and in particular of HIV. It has been found that reduction of the transmissibility without any information based on the connectivity of the network is not a good strategy. The more oriented a strategy is towards the nodes with large *k*, the more chance it has to bring the epidemic threshold below the disease spreading rate.

References:

Bauch, C. and Rand, D. A. (2000). *A moment closure model for sexually transmitted disease transmission through a concurrent partnership network.* Proc. R. Soc. London. B., 267:2019–27.

Diekmann, O. and Heesterbeek, J. A. P. (2000). *Mathematical Epidemiology of Infectious Diseases.* John Wiley & Son, Chichester.

Diekmann, O., De-Jong, M. C. and Metz, J. A. (1998). *A deterministic epidemic model taking account of repeated contacts between the same individuals.* J. Appl. Prob., 35:448–62.

Eguiluz, V. M. and Klemm, K. (2002). *Epidemic threshold in structured scale-free networks*. Phys. Rev. Lett., 89:108701.

Ferguson, N. M. and Garnett, G. P. (2000). *More realistic models of sexually transmitted disease transmission dynamics -sexual partnership networks, pair models, and moment closure.* Sex. Trans. Dis., 27:600–9.

Ferguson, N. M. , Donnelly, C. A. and Anderson, R. M. (2001). *Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain*. Nature, 413:542–8.

Ghani, A. C. and Garnett, G. P. (1998). *Measuring sexual partner networks for transmission of sexually transmitted diseases*. J. R. Stat. Soc. A, 161:227–38.

Ghani, A. C. and Garnett, G. P. (2000). *Risks of acquiring and transmitting sexually transmitted diseases in sexual partner networks.* Sex. Trans. Dis., 27:579–87.

Ghani, A. C., Donnelly, C. A. and Garnett G. P. (1998). *Sampling biases and missing data in explorations of sexual partner networks for the spread of sexually transmitted diseases*. Stat. Med., 17:2079–97.

Hethcote, H. W. and Van Ark, J. W. (1987). *Epidemiological models for heterogeneous populations: proportionate mixing, parameter estimation, and immunization programs.* Math. Biosci., 84:85–118.

Jones, J. H. and Handcock, M. S. (2003). Social networks: *Sexual contacts and epidemic thresholds*. Nature, 423:605–6.

Keeling, M. J. (1999). *Correlation equations for endemic diseases: externally imposed and internally generated heterogeneity*. Proc. R. Soc. Lond. B, 266:953–60.

Latora, V., Andre N., Jacques S., Bahire S., Sandwidi D., Bukiki S., and Salvatore M. (2006).*Network of Sexual Contacts and Sexually Transmitted HIV Infection in Burkina Faso*, Journal of Medical Virology 78:724–729

Liljeros, F. ,Edling, C. R. , Amaral, L. A., Stanley, H. E. and Aberg, Y. (2001). *The web of human sexual contacts.* Nature, 411:907–8.

Meyers, L. A, Pourbohloul, B, Newman, M. E., Skowronski, D. M., and Brunham, R. C. (2005). *Network theory and SARS: predicting outbreak diversity*. J. Theor. Biol., 232:71–81.

Morris, M. and Kretzschmar, M. (1997). *Concurrent partnerships and the spread of HIV. AIDS*, 11:641–8.

Pastor-Satorras, R. and Vespignani, A.(2001). *Epidemic spreading in scale-free networks.* Phys. Rev. Lett., 86:3200–3.

Pastor-Satorras, R. and Vespignani, A.(2002). *Immunization of complex networks*. Phys. Rev. E, 65:036104.