



# Sexual networks in contemporary Western societies

Fredrik Liljeros<sup>a,b,\*</sup>

<sup>a</sup>*Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Solna, Sweden*

<sup>b</sup>*Department of Sociology, Stockholm University, Stockholm, Sweden*

---

## Abstract

Sexually transmitted infections continue to be a severe health problem in contemporary Western societies, despite the considerable funds allocated for control programs. In this article, we present and discuss a variety of explanations that have been advanced on why this type of disease is so hard to eradicate, despite the fact that the contact by which it is spread is far less frequent than is the case with most other infectious diseases. We conclude that several processes and mechanisms facilitate the spread of sexually infected diseases, and that both broad and targeted intervention is therefore needed to eradicate such diseases.

© 2004 Elsevier B.V. All rights reserved.

*PACS:* 87.23.Ge; 87.19.Xx

*Keywords:* Social networks; Infectious diseases

---

## 1. Introduction

In spite of control efforts such as targeted screening, contact tracing, and information campaigns, sexually transmitted infections (STIs) such as chlamydia, gonorrhea, and HIV are still a problem in Western societies. The fact that STIs remain endemic in Western societies despite the low average annual number of sexual partners reported in national surveys in these societies is not yet fully understood [1–3]. In this paper, we review and discuss some of the explanatory mechanisms that have been proposed.

---

\* Corresponding author. Fax: +46 8 612 55 80.

*E-mail address:* [liljeros@sociology.su.se](mailto:liljeros@sociology.su.se) (F. Liljeros).

## 2. Data sources

The study of sexual networks can to some extent be compared to the study of planets in other solar systems in the sense that they can only be studied indirectly. Our knowledge comes from at least four different sources, each of which has its specific advantages and disadvantages. The first consists of national surveys of sexual behavior [1–3]. Such surveys have been carried out in most Western societies. The advantage of such studies is that, given that they are based on random population samples, they are the only kind of study that can theoretically give us knowledge about sexual behavior for the whole population. Unfortunately, such studies have several drawbacks. It is very hard to obtain the accuracy of respondents' answers. The response rate is also usually too low to guarantee reliable precision in the estimates.

One striking thing about national surveys, for example, is that men on average report a significantly higher number of sexual partners than do women. This discrepancy has been explained by a tendency for men as a group to over-report the number of sexual partners [4]. A recent study has also shown that this difference may be explained by the fact that prostitutes are usually not included in the samples [5]. National surveys have another disadvantage that is probably more important than the validity problem mentioned above, namely, that they are only able to give us information about the behavior of the respondents, and not information about the behavior of their sexual partners (and *their* partners). National surveys can therefore not give us information about the global properties of a network, such as level of clustering, size of the largest interconnected component, or average distance between the individuals.

A second source of information about the sexual network is the network data generated by contact tracing [6]. Contact tracing is the process whereby the contacts of an individual who has tested positive for an STI are traced and tested. If this procedure is also continued for the contacts that tested positive, and for their contacts in turn, it is eventually possible to generate a sub-graph of the sexual network. This sub-network can then be used to analyze global structural properties of the network that cannot be studied with national survey data. This kind of sampling is also associated with severe biases. It is not always possible for an individual, even if she is cooperative, to give enough information about any given sexual partner to be able to identify him or her. A more serious bias is that by definition contact tracing has a tendency to identify the sub-networks of the general sexual networks in which it is easiest for the disease to spread. The latter may, however, not always be a problem. If, for example, the purpose of a study is to find ways to mitigate the epidemic in the parts of the network in which the STIs are spread, this type of biased data really can be very useful.

A third type of data source is the mapping of the sexual network. This has so far only been carried out at the local level, probably for practical reasons, for example with a group of drug users and prostitutes in a US town [7] and in a high school in the United States [8]. These studies have the same problem as contact tracing studies when it comes to identifying and finding partners. It is also important to remember that such studies cannot be generalized to the rest of the population, as studies based on a random sample can.

The fourth and last data source we are going to mention here is the data generated from Internet dating communities [9]. Because much of the activity that takes place on such communities is logged, it is possible to extract a network about which members interact. The advantage with this data source is that all contacts in the community can be traced, and that networks of a significant size can be generated. This data source is, however, probably the least valid network discussed here when it comes to mapping the actual sexual network, because no information exists about actual sexual contact.

The classification of data sources into the four groups cited here should be seen as a preliminary one. There are also combinations of them such as surveys based on non-random samples, for example, individuals who are seen at STI clinics [10], or convenient samples of university students. This short survey of data sources for sexual networks shows that all data sources are associated with different kinds of validity problems. This is a quite common situation in the social sciences. One way, if not to solve but at least to mitigate the problem, is to try to confirm an empirical result with data from different sources, a procedure usually referred to as triangulation [11].

### 3. The standard model of infectious diseases

In standard epidemiological models, persons are assumed to be in either one of three states. According to these states, the population can be classified into one of the three categories: susceptible (S), infected (I), or resistant (R) persons. It is conventional to distinguish between SI, SIS, and SIR models, and in all of these, interaction between persons is assumed to be random and homogeneous [12]. For most sexually transmitted diseases, the SIS model makes the most sense since few sexually transmitted diseases confer any immunity after infection. An important exception is HIV, which is still appropriately described with the SI model.

Random interaction is assumed in Eq. (2), which gives the SIS model as a continuous model in its simplest form, consisting of a system of two differential equations,

$$\begin{aligned} \frac{dS}{dt} &= \frac{-c\beta S(t)I(t)}{N} + \frac{I(t)}{D}, \\ \frac{dI}{dt} &= \frac{c\beta S(t)I(t)}{N} - \frac{I(t)}{D}. \end{aligned} \quad (1)$$

There are two-dependent variables in Eq. (2): The number of susceptible persons,  $S$ , and the number of infected persons,  $I$ . At each point in time  $S(t) + I(t) = N$ , where  $N$  is the population size, parameter  $c$  is the number of potentially infectious contacts per person, parameter  $\beta$  is the probability of infection per contact between susceptible and infected, and  $D$  is the mean duration of infection [13]. Evidently, this model is homogenous across persons as each person is assumed to have the same number of contacts,  $c$ . Of course, in the case of STI, this simple model would also take into account gender differences, etc. by adding more equations to the system of equations [14,15]. A significant advantage of the random interaction assumption is that it can easily be modeled with differential equations, and these models can be studied analytically [16].

A critical notion in disease epidemiology is the basic reproduction number,  $R_0$ . In the homogenous deterministic model (Eq. (2)) this number tells us as to how many uninfected persons an infected person will, on average, transmit his or her infection in a totally susceptible population [12]

$$R_0 = c\beta D . \quad (2)$$

$R_0$  has received special attention because in the homogenous model it is quite intuitive. If  $R_0$  is less than 1, then the disease will become extinct. With  $R_0$  equaling exactly one, we have an unstable equilibrium with no change in the number of infected or susceptible persons, that is, the disease is endemic. A value greater than 1 upsets the replacement situation, which means that if  $R_0 > 1$  the outcome is an epidemic.

#### 4. Sexual contact patterns

Several explanations have been suggested on why STIs are still endemic in Western societies. All of them are related to specific structural properties of the sexual contact patterns in these societies. The core group theory is an early, and probably also the best known, explanation for the endemic situation [17]. According to the core group theory, the reproduction of STIs can be explained by the existence of several distinct sub-groups in the general population that are all characterized by high-risk sexual behavior and extensive inter-group interaction. A tendency for individuals to prefer sexual contact with persons similar to themselves is usually referred to as *assortative interaction* in the STI literature. A tendency to assortative interaction has been reported for social factors such as social class and ethnicity [18]. The existence of core groups, according to the theory, makes it possible for STIs to reproduce themselves within the core group because the  $R_0$  is greater than 1. The core groups thus constitute the reservoir needed for the STI to remain endemic in a general population in which the  $R_0$  is lower than the critical value 1. The core groups also make it possible for the rest of the population to be infected through contacts with this group.

In the first models of populations containing core groups, the difference between individuals in the core groups and the rest of the population was assumed to be merely a difference in contact rate. Later research has yielded a more elaborated picture of the characteristic feature of the core groups. The emergence of HIV in the early 1980s generated intense research activity into sexual contact patterns and their importance for the spread of HIV [19]. One of the most important findings was made by Anderson and May [12], who showed that the expression for the  $R_0$  in expression 2 is not suitable if the variance in the number of potential infectious contacts is high, as is the case with STIs. They showed instead that the  $R_0$  would be more accurately estimated by using the following expression for  $R_0$ :

$$R_0 = \rho_0 \left( 1 + \frac{\sigma^2}{\mu} \right) , \quad (3)$$

where  $\rho_0$  is the average number of infections produced by an infected person in an uninfected population,  $\sigma^2$  is the variance in the number of contacts, and  $\mu$  is the mean number of contacts in the population. From this equation, it is clear that the larger the

variance in the number of partners in the population for a given  $\mu$ , the less infectious an infection needs to be to continue to reproduce itself, that is, to generate an epidemic. It is also interesting that this finding neither presupposes a clearly defined set of core groups nor does it presuppose a general tendency to assortative interaction greater than the natural consequence of the distribution itself.

The traditional epidemiological model based on homogeneous mixing generally predicts exponential growth in the number of infected persons during the early stages of an epidemic. This was, however, not in line with the observation of the growth of HIV infection in the 1980s. The cumulative growth of the new disease showed a slower than exponential growth and could better be described by a cubic growth over time [20]. This anomaly inspired Colgate et al. [10] to suggest a model for the spread of HIV in which the time an individual became infected is negatively related to the number of partners the individuals has per time unit, that is, individuals with a large number of partners are likely to be infected during the early stages of the outbreak of the disease. An interesting feature of their model is that both a strong assumption about assortative interaction and a highly skewed power law distribution of the number of sexual contacts is needed for the model to generate a growth pattern similar to what was observed empirically. The authors were also able to confirm their prediction of a power law distribution in the number of contacts on clinical data of a population of gay men seen at an STI clinic in London.

It is well known in the social sciences that an empirical observation made in one group of people does not necessarily hold true in other social settings. Interestingly, a skewed power law distribution has recently also been reported for a Swedish population, both for numbers of partners during last 12 month and life time number of partners [21]. Whether power law tails in the upper tail of the frequency distributions of sexual contacts are a universal feature of human behavior is, however, still a disputed question [22–24].

A power law distribution in the number of sexual contacts has, under some specific circumstances, recently been shown to have serious consequences for the potential to eradicate STIs [25,26]. In an infinite population with homogeneous mixing, a slope smaller than three makes the second moment of the distribution infinitely large, and therefore also the variance of the distribution [27]. It can easily be seen from Eq. (3) that this will result in an infinitely large  $R_0$  for the population. This results in the bizarre consequence that all individuals must be tested and treated at the same time, otherwise it will be impossible to eradicate the disease. There are, however, several reasons to believe that this is not the case. We know, for example, that all human populations must by definition have a finite size. Even though it has been shown that a power law distribution in the number of contacts will increase the  $R_0$  significantly in a finite population compared to a population with the same mean number of contacts but a low variance in number of contacts (which also follows from Eq. (4)) [28]; this does not imply that an epidemic outbreak of an STI cannot be stopped or at least curtailed. Another factor that probably mitigates the effect of the skewed distribution is that there must also be an upper limit for the number of sexual contacts a single individual can have per unit of time. A study of prostitutes in the United States shows, however, that this limit can be very large. The median value for number of partners

was found to be as large as 103 during the previous 6 months [5]. One advantage of the skewed distribution, from an STI prevention perspective, is that it is predicted that the  $R_0$  can be drastically reduced and the epidemic eventually stopped if the individuals who change partners frequently can be tested and convinced to practice safe sex [29]. It may at a first glance, seem difficult to identify individuals who change partners frequently for specific targeted intervention, except perhaps for specific groups such as prostitutes, or gay men who visit video clubs where anonymous sex takes place. Contact tracing, however, has the positive side effect that individuals who have many contacts have a higher probability of showing up because they have a larger groups of sources of infection from than do individuals with fewer contacts [30].

The number of unprotected sexual contacts is clearly associated with an individuals risk of STI. It is, however, possible that the number of sexual contacts is not the most important risk factor for getting an STI. Morris and Kretzschmar suggested in a series of articles [31–33] that it is the frequency of concurrent relations (partnerings that overlap over short time periods) in a population that is the most important factor in the transmission of STIs. Concurrent relations are important because potential contacts for the transmission of an STI come much closer in time if individuals have concurrent relations than if they practiced serial monogamy.

The spatial diffusion of STIs over greater geographical distance has recently received increased attention in the STI research community, partly because of the current potential for identifying different strains of STIs through genotyping [6,34,35]. Watts and Strogatz have, in their now classical 1998 paper on small-world networks, showed that only a small fraction of random links is needed to decrease drastically the mean distance in a spatially structured network [36]. It is interesting in this context that a study by Falk et al. [37] reported that 22 individuals tested positive for chlamydia at a hospital in a medium sized Swedish town were infected abroad in 11 different countries on four different continents. (They also reported strong evidence for an indirect link to a fifth continent). This study and many others clearly indicate that the web of sexual contacts that connects different towns and geographical areas is very likely a small-world network characterized by a large amount of local interaction, but at the same time by a significant number of contacts over longer distances. (Sexual networks of heterosexual individuals can by definition not be a small-world network as defined by Watts and Strogatz because the level of clustering must by definition be zero in these networks as no network cycles can exist that consist of three units.)

According to a working paper that is circulating in the STI community, we may expect many interesting papers in the near future. Two studies are specially worth mentioning. The first one is a recent study of romantic and sexual relationship in a high school in the United States that reports what looks like a strong tendency among the individuals not to form network cycles of four. We must, as always, be cautious when making inferences from one specific local setting to a general population, but if this norm also exists in other social settings, could this be a factor that will make it easier for a giant interconnected component to be formed than we might otherwise expect? The second study is reported in a recent working paper [38]. Based on a simulation model it suggests that the small set of individuals that reports a large number of partners in sexual surveys are not necessarily as important for transmitting STIs as

generally thought. The authors argue instead that multiconnected components, that is, network components such as network cycles in which each individual reaches each other in mutually exclusive ways, may be of greater importance for the spread of STIs. Interestingly, their preliminary simulation results show that large bicomponents, that is, components in which each individual can reach each other individual in two different, mutually exclusive ways, may emerge in networks that have a low variance and a relatively low mean number of contacts.

The results presented here indicate that many different structural properties must be taken into consideration to understand the spread of STIs in contemporary Western societies. It is therefore not likely that one single solution to the problem ever will be found. The solution probably lies in a combination of broad and targeted interventions. It is also important to remember that control efforts that have been made in many Western societies have already drastically decreased the incidence of STIs, especially gonorrhoea [39] (even though the disease has not been totally eradicated). The fact that an increase in the incidence of STIs has been reported in some Western countries is a warning that must be taken seriously [39,40]. It is also important to remember that STIs are now a global problem, and as long as travel patterns between countries and continents persist, STIs will never be eradicated in the Western world as long as STIs is still endemic in the rest of the world. The study mentioned above from the small Swedish town [37] shows very clearly that the sexual network is a global network.

A full understanding of how STIs are spread in Western societies is probably not possible unless they are transformed into Orwellian societies in which all sexual activity is monitored. This brief survey shows, however, that much insight can be gained with the data we have in hand. The field of sexual networks has become an intense research area that attracts researchers from many different fields [18]. This fact makes it reasonable to think that much progress can be made in the future when it comes to increased understanding of the spread of STIs in Western societies and to developing effective methods for controlling them. One current problem is that the progress that has been made in the different fields is not always communicated. This can probably be explained to some extent by the fact that there is little tradition among leading researchers in the field for writing regular, comprehensive review articles about progress in the field. Perhaps this brief introduction can inspire such articles.

## Acknowledgements

I would like to thank L.A.N. Amaral and C.R. Edling for helpful comments on earlier versions of this manuscript. Financial support from The National Institute of Public Health (NIPH), Sweden is acknowledged.

## References

- [1] B. Lewin (Ed.), *Sex in Sweden—on the Swedish Sexual Life*, The National Institute of Public Health in Sweden, Stockholm, 2000.
- [2] A.M. Johnson, et al., *Lancet* 358 (2001) 1835–1842.

- [3] E.O. Laumann, Y. Youm, *Sexually Transmitted Dis.* 26 (1999) 250–261.
- [4] M. Morris, *Nature* 365 (1993) 437–440.
- [5] D.D. Brewer, et al., *Proceedings of the National Academy of Sciences of the United States of America* 97 (2000) 12385–12388.
- [6] J.L. Wylie, A. Jolly, *Sexually Transmitted Dis.* 28 (2001) 14–24.
- [7] R.B. Rothenberg, et al., *Aids* 14 (2000) 2191–2200.
- [8] P.S. Bearman, J. Moody, K. Stovel, *Am. J. Sociol.*, 2002, accepted for publication.
- [9] P. Holme, C.R. Edling, F. Liljeros, *Social Networks*, in press.
- [10] S.A. Colgate, E.A. Stanley, J.M. Hyman, S.P. Layne, C. Qualls, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 86, 1989, pp. 4793–4797.
- [11] C.F. Nachmias, D. Nachmias, *Research Methods in The Social Sciences*, Eward Arnold, London, 1996.
- [12] P.R. Anderson, R.M. May, *Infectious Diseases of Humans*, Oxford University Press, Oxford, 1991.
- [13] J. Giesecke, *Modern Infectious Disease Epidemiology*, Eward Arnold, London, 1994.
- [14] M. Altmann, *J. Math. Biol.* 33 (1995) 661–675.
- [15] M. Altmann, *Math. Biosci.* 150 (1998) 153–175.
- [16] O. Diekman, J.A.P. Heesterbeek, *Mathematical Epidemiology of Infectious Disease*, Wiley, Chichester, 2000.
- [17] H. Hethcote J.A. Yorke, *Gonorrhea Transmission Dynamics and Control*, Springer, New York, 1984.
- [18] F. Liljeros, C.R. Edling, L.A.N. Amaral, *Microbes Infect.* 2003 (2003) 189–196.
- [19] A.S. Klondahl, *Social Sci. Med.* 21 (1985) 1203–1216.
- [20] J.M. Hyman, E.A. Stanley, *Math. Biosci.* 90 (1988) 415–473.
- [21] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley, Y. Aberg, *Nature* 411 (2001) 907–908.
- [22] F. Liljeros, C.R. Edling, H.E. Stanley, Y. Aberg, L.A.N. Amaral, *Nature* 423 (2003) 606.
- [23] J.H. Jones, M.S. Handcock, *Nature* 423 (2003) 605–606.
- [24] J.H. Jones, M.S. Handcock, *Proceedings of the Royal Society of London Series B-Biological Sciences* 270 (2003) 1123–1128.
- [25] R. Pastor-Satorras, A. Vespignani, *Phys. Rev. Lett.* 86 (2001) 3200–3203.
- [26] R. Pastor-Satorras, A. Vespignani, *Phys. Rev. E* 63 (2001) 066117.
- [27] A.L. Lloyd, R.M. May, *Science* 292 (2001) 1316–1317.
- [28] R. Pastor-Satorras, A. Vespignani, *Phys. Rev. E* 65 (2002) (art. no.-035108).
- [29] Z. Dezso, A.L. Barabasi, *Phys. Rev. E* 65 (2002) (art. no.-055103).
- [30] M.E.J. Newman, *Social Networks* 25 (2003) 83–95.
- [31] M. Kretzschmar, M. Morris, *Math. Biosci.* 133 (1996) 165–195.
- [32] M. Morris, M. Kretzschmar, *Social Networks* 17 (1995) 299–318.
- [33] M. Morris, M. Kretzschmar, *Aids* 11 (1997) 641–648.
- [34] T. Cabral, A.M. Jolly, J.L. Wylie, *J. Infect. Dis.* 187 (2003) 279–286.
- [35] R.P. Kerani, et al., *Sexually Transmitted Dis.* 30 (2003) 742–749.
- [36] D.J. Watts, S.H. Strogatz, *Nature* 393 (1998) 440–442.
- [37] L. Falk, et al., *Sexually Transmitted Dis.* 30 (2003) 205–210.
- [38] J. Moody, M. Morris, J. Adams, M.S. Handcock, *The Ohio State University*, Columbus, 2003.
- [39] Report, *Swedish Institute for Infectious Disease Control*, Solna, 2003.
- [40] E. Hiltunen-Back, et al., *Sexually Transmitted Dis.* 30 (2003) 737–741.