



HAL
open science

Weighting for sex acts to understand the spread of STI on networks

Mathieu Moslonka-Lefebvre, Sebastian Bonhoeffer, Samuel Alizon

► **To cite this version:**

Mathieu Moslonka-Lefebvre, Sebastian Bonhoeffer, Samuel Alizon. Weighting for sex acts to understand the spread of STI on networks. *Journal of Theoretical Biology*, 2012, 311 (octobre 2012 et janvier 2013), pp.46-53. 10.1016/j.jtbi.2012.06.031 . hal-01567925

HAL Id: hal-01567925

<https://hal.science/hal-01567925>

Submitted on 28 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Weighting for sex acts to understand the spread of STI on networks

Mathieu Moslonka-Lefebvre^{a,b,*}, Sebastian Bonhoeffer^a, Samuel Alizon^c

^a Institute for Integrative Biology, ETH, Universitätstrasse 16, 8092 Zürich, Switzerland

^b INRA, UR 341 Mathématiques et Informatique Appliquées, 78350 Jouy-en-Josas, France

^c Laboratoire MIVEGEC (UMR CNRS 5290, IRD 224, UM1, UM2), 911 Avenue Agropolis, 34394 Montpellier Cedex 5, France

HIGHLIGHTS

- ▶ Human sexual networks exhibit a heterogeneous structure.
- ▶ Expected network-based properties are at odds with epidemiological data for STI.
- ▶ Most network models typically assume a constant transmission risk per partnership.
- ▶ We develop a framework to weight sexual networks based on biological assumptions.
- ▶ Our weighting model re-conciliate network theory and epidemiological data for STI.

ARTICLE INFO

Article history:

Received 20 February 2012

Received in revised form

21 June 2012

Accepted 25 June 2012

Available online 3 July 2012

Keywords:

Weighted networks

Epidemiology

STI

Heterogeneity

Sex acts

ABSTRACT

Human sexual networks exhibit a heterogeneous structure where few individuals have many partners and many individuals have few partners. Network theory predicts that the spread of sexually transmitted infections (STI) on such networks should exhibit striking properties (e.g. rapid spread). However, these properties cannot be found in epidemiological data. Current network models typically assume a constant STI transmission risk per partnership, which is unrealistic because it implies that sexual activity is proportional to the number of partners and that individuals have the same activity with each partner. We develop a framework that allows us to weight any sexual network based on biological assumptions. Our results indicate that STI spreading on the resulting weighted networks do not have heterogeneous-related properties, which is consistent with data and earlier studies.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Networks reflecting human sexual contacts over several years in a population have been shown to exhibit a high variance in the number of sexual partners per individual (Liljeros et al., 2001; Hamilton et al., 2008). The exact nature of the distribution of partners per individual is debated: some argue that it tends to follow a power-law (PL; Liljeros et al., 2001; Schneeberger et al., 2004; Hamilton et al., 2008), others argue that it tends to follow a negative-binomial (NB; Handcock and Jones, 2004; Hamilton et al., 2008). At any rate, both these distributions are the appanage of highly heterogeneous networks in which a few nodes in a network (i.e. individuals) are highly connected (i.e. have many partners), while the others nodes have few connections. Theory predicts that epidemics spreading on heterogeneous

networks (either PL or NB) should exhibit specific properties such as a very low epidemic threshold, a rapid spread or in the case of PL, a doubling time that decreases with population size (Newman, 2003; Keeling and Eames, 2005). There is an ongoing debate in the literature concerning the relevance of these predictions for biological systems (especially the absence of epidemic threshold; May, 2006; Hamilton et al., 2008). For instance, specific epidemiological properties of highly heterogeneous networks are not observed generally in the data (Liljeros, 2004; Handcock and Jones, 2006; Britton et al., 2007; Hamilton et al., 2008).

Most network-based epidemiological studies consider unweighted networks. In the context of sexually transmitted infections (STI), it means that the risk of infection is the same on all the edges (i.e. for all the interactions) in the network (but see Eames et al., 2009). While these assumptions may be justifiable for diseases such as flu or tuberculosis that only require short-term interactions between hosts, they are unrealistic for STI. Here, we develop a new framework to add sex acts explicitly on sexual contact networks and show that such a biologically relevant weighting has major epidemiological consequences.

* Corresponding author at: INRA, UR 341 Mathématiques et Informatique Appliquées, 78350 Jouy-en-Josas, France. Tel.: +33 1 34 65 22 47.

E-mail address: mathieu.moslonka-lefebvre@jouy.inra.fr
(M. Moslonka-Lefebvre).

Modelling the spread of STI on unweighted networks makes two implicit assumptions. First, the total number of sex acts of an individual per unit of time is assumed to be strictly proportional to his/her number of sexual partners. Second, these sex acts are assumed to be partitioned equally among all the partners. That individuals with five sexual partners generally have five times more sex acts per unit of time than individuals with one partner is unrealistic and contradicts empirical data (Blower and Boe, 1993; Nordvik and Liljeros, 2006; Britton et al., 2007).

Several studies point out the importance of weighted networks in general (see e.g. Barrat et al., 2004b's seminal work). Some studies consider the spread of STIs on weighted networks of infinite size. Newman (2002) already showed that assuming that the transmission rate from one infected individual to another is not constant but rather a function of node degree (i.e. number of partners) can affect disease spread. More precisely, he shows that for epidemics to spread rates of transmission need to fall off slower than inversely with node degree (independently of the network structure). Wang et al. (2007) approximate an infinite size so-called 'scale free' network with mean field equations and show that assuming that the transmission from one node to another is proportional to the node degree affects epidemic spreading. Joo and Lebowitz (2004) use a similar approach but consider a more elaborate transmission function, which they allow to saturate with increasing node degree. They find that such a saturation further decreases the speed of disease spread. Recently, Britton et al. (2011) study infinite size weighted sexual contact networks, where weights are drawn in a distribution that is allowed to depend on node degree. They show that if nodes (vertices) with high degree tend to have a low weight, then it is harder for an epidemics to take off. The two main limitations of these models are that they consider infinite size networks and/or that they do not consider any biological basis for the weighting they use. Note also that in most of these formalisms, the transmission rate from an individual A to an individual B differs from the transmission rate from B to A if they have different number of partners. In Britton et al. (2011) the assumption is slightly different because a node can only be linked to a node that has the same weight (which means the weighting constraints the shape of the network).

More biologically oriented approaches develop individual-based models to understand how sexual contact networks emerge (see e.g. Althaus et al., 2010) or pair-approximation models but without taking into account sex acts (Ferguson and Garnett, 2000). Among these studies, two consider the effect of sexual activity on the spread of an STI more explicitly. Röttingen and Garnett (2002) study the association between STI risk and number of sex acts without a network-based approach by fitting the relationship between the HIV transmission probability per-partnership and the number of sex acts. However, their model ignores epidemiological feedbacks, thus assuming for instance that all sexual partners of one individual have the same probability to be infected (whereas this probability should at least depend on the number of partner these partners have). More recently, Britton et al. (2007) model the spread of an STI in a heterogeneous population (in terms of number of sexual partners) where individual of each class can have both 'steady' or 'casual' partnerships. The latter partnerships have higher transmission probabilities than the former and the ratio of each type of partnership is inferred from empirical data on the number of sex acts per partner. They show that the basic reproduction number (R_0) can be over-estimated when different types of partnerships are not considered.

For sexual contact networks, the weight of an edge between two nodes corresponds to the number of sex acts that are actually realised between two individuals per unit of time (here weeks).

In the following, we refer to these as realised sex acts (RSA). We derive the number of RSA between any interacting pair (i.e. all the weights of the network edges) from the network topology based on two simple biological assumptions. First, each individual is allocated a number of potential sex acts (PSA). PSA can be seen as a quantification of the sexual activity of an individual. Second, each individual partitions his/her PSA among his/her sexual partners. The number of RSA between two individuals i and j depends on the number of PSA that i attributed to j and vice versa (see the Methods section).

We consider three models that describe the link between the number of PSA allocated to an individual and the number of partners of this individual: a linear relationship (the linear allocation model A_{lin}), a saturating relationship (the saturating allocation model A_{sat}), or a constant relationship where each individual receive the same number of PSA (the constant allocation model, A_{cst}). Further details about the model are available in the Material and Methods section. Classical models implicitly make the strong assumption that PSA are allocated linearly with the number of sex partners, which is at odds with empirical data (Blower and Boe, 1993; Nordvik and Liljeros, 2006).

In our study, PSA can then be partitioned by an individual among his/her partners in three ways: (i) equally (P_{equ}), where all the partners get the same share, (ii) randomly (P_{rand}), where the fraction each partner gets is random and (iii) to maximise his/her number of RSA (P_{max}). Classical models implicitly assume an equal or a maximising partitioning, but the random partitioning arguably better reflects current sexuality, where the majority does not act in terms of maximising sex budgets (Foucault, 1976).

Our framework stands out compared to earlier studies because we use a finite size network inferred from contact tracing data, contrary to earlier models that are based on infinite 'scale-free' networks (Joo and Lebowitz, 2004; Wang et al., 2007) or that use degree distributions without explicit network topology (Röttingen and Garnett, 2002; Britton et al., 2007). This allows us to model explicitly how individual sexual behaviours affect the network weighted topology and, hence disease spread. Earlier studies that account for sex acts either do not include epidemiological feedbacks (Röttingen and Garnett, 2002) or only model a heterogenous population with two types of partnerships and therefore cannot keep track of individual behaviours (Britton et al., 2007). Our framework also allows us to model sex acts explicitly on any network type. Overall, we show that minimal deviations from implicit assumptions regarding sexual behaviour made in earlier models can strongly affect diseases spread.

2. The model

2.1. The network

A sexual network is a representation of individuals (the network nodes) and sexual relations between these individuals (the network edges). Each individual i has k_i sexual partners. Sexual networks evolve through time by creation and removal of sexual relationships among partners. However, the networks we usually have access to summarise sexual contacts over long periods of time because, in most studies, participants are asked to list their sexual partners over several years (see e.g. Potterat et al., 1985; Liljeros, 2004). Our method of weighting allows the modelling of epidemics occurring on a short time scale (days), even with a contact network inferred over a longer time scale (years).

For simplicity, we focus on disease spread on static networks of men having sex with men (MSM), even if our methods are also applicable to dynamical and bipartite (heterosexual) networks. Such networks can be described using an adjacency matrix (A),

the elements of which ($\mathcal{A}(i,j)$) are equal to 1 if there is a link connecting node i to node j and 0 otherwise.

Except when stated otherwise, the results we present are obtained on an empirical network of 381 nodes built from a contact tracing study in Colorado Springs involving a cohort of individuals infected with gonorrhoea (Potterat et al., 1985). The average degree of the network (i.e. number of sexual partners) is 2. As shown in the Supplementary Methods, the degree distribution of this network is best described by a power law of coefficient $\gamma = 1.98$.

We also generated theoretical power-law (PL) and negative-binomial (NB) networks of different sizes with the Barabasi–Albert algorithm with a γ coefficient of 3 (Barabasi and Albert, 1999) and the so-called ‘configuration model’ (Catanzaro et al., 2005) respectively. For both PL and NB, we set the average degree to 4 (see Supplementary Methods). This allowed us to check that our results were not due to the specific topology of the empirical network.

2.2. Weighting the network

Each individual i is allocated a number of potential sex acts (PSA) per unit of time (denoted b_i). This individual then partitions his/her PSA among his/her k_i sexual partners. It is thus possible to define the availability of an individual i for interacting with j , which is the number of PSA allocated by i to j (denoted $[i \rightarrow j]$).

2.2.1. Allocation of sex acts

We consider three models as to how the number of PSA (b_i) is allocated to an individual i :

- The linear allocation model (A_{lin}), which assumes that b_i is strictly proportional to k_i , the number of partners.
- The saturating allocation model (A_{sat}), which assumes that b_i is proportional to $\sqrt{k_i}$.
- The constant allocation model (A_{cst}), which assumes that b_i is the same for any node of the network.

2.2.2. Partitioning of sex acts

PSA are then partitioned by an individual i among his/her partners according to three models:

- The equal partitioning model (P_{equ}), where all the partners of an individual i get an equal share $[i \rightarrow j] = b_i/k_i$.
- The random partitioning model (P_{rand}), where all the partners of an individual i get a share $[i \rightarrow j] = b_i \rho_j/R$, where ρ_j is drawn in a uniform distribution and $R = \sum_{j=1}^{k_i} \rho_j$.
- The maximisation partitioning model (P_{max}), in which an individual uses information about the topology to maximise their number of realised sex acts by giving to his/her partners in proportion to what these partners are giving to him (see the Supplementary Methods for further details).

STI models with an unweighted network implicitly assume a linear allocation model with an equal or a maximising partitioning.

2.2.3. Realised sex acts

The number of sex acts actually realised (RSA) between individuals i and j (denoted $W(i,j)$) is then assumed to be

$$W(i,j) = \min([i \rightarrow j], [j \rightarrow i]) \quad (1)$$

because an individual i cannot have more sex acts with a partner j than the number of sex acts that j is willing to have with i and vice versa. From this, we obviously have $W(i,j) = W(j,i)$.

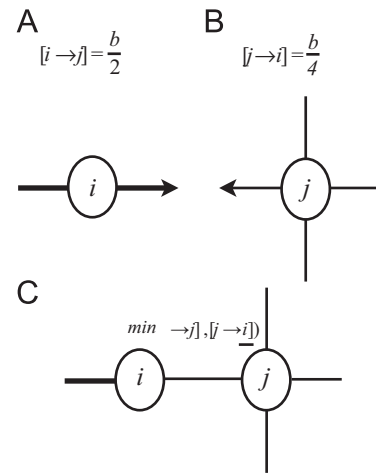


Fig. 1. Illustration of our weighting model. In this example, individuals i (A) and j (B) are allocated a constant sex budget (set to $b=4$ sex acts per week) that they partition equally among their partners. The number of sex acts actually realised (RSA) between individuals i and j is the minimum of the two contributions (C).

Our weighting framework (Fig. 1) thus allows us to transform a classical adjacency matrix (\mathcal{A}) into a weighted adjacency matrix (W), the elements of which are the realised number of sex acts (as defined in Eq. (1)).

2.2.4. Loss of sex acts due to the weighting

Care is needed because, even though our definition on how sex acts are realised is biologically sensible, it introduces a potential confounding factor because the more $[i \rightarrow j]$ and $[j \rightarrow i]$ differ, the more PSA will be lost. We quantify this loss of sex acts on the whole network with a parameter δ defined as

$$\delta = \frac{\sum_i \sum_j W(i,j)}{\sum_i \sum_j s^* \mathcal{A}(i,j)} \quad (2)$$

where the $W(i,j)$ are the elements of the weighted adjacency matrix, $\mathcal{A}(i,j)$ are the elements of the adjacency matrix and s^* is the average number of potential sex acts between two individuals.

We want to compare the spread of a pathogen on different weighted networks and this spread is obviously a function of the total number of sex acts realised on the network. In order to isolate the effect of the network itself, we define the weights on the network such that the average number of sex acts realised per individual per unit of time on a network is equal to the average degree of this network ($s^* = \bar{k} = 4$). We hence calculate a corrected weighted adjacency matrix $\tilde{W} = W/\delta$. Note that for the genuine GonoCocci network, denoted the GC network throughout, $\bar{k}_{GC} = 2$ so in order to get values comparable to the other networks we corrected the W for GC by $2\bar{k}_{GC}$. In summary, the total number of realised sex acts is constant for any network structure, allocation model and partitioning model.

2.3. The epidemiological model

Since our goal is to investigate the importance of weighting the network, we deliberately consider a simple epidemiological SI model, where hosts can be either susceptible or infected (Anderson and May, 1991). When an infected individual has sex with a susceptible one, the disease is assumed to be transmitted at a constant rate per sex act. We further assume that there is no recovery and that death does not occur over the duration of the simulation.

This simple model is a reasonably good approximation to the transmission of HIV over short periods of time (Liljeros, 2004), as there is no significant removal of individuals due to death on timescales close to the average duration of an HIV infection if untreated (≈ 10 year; Anderson and May, 1991). Furthermore, transmission rate per sex acts was set to $\beta = 0.01$, which is consistent with HIV data (Vittinghoff et al., 1999; Boily et al., 2009). The speed of disease spread is estimated through the median doubling time t_d . The doubling time is an appropriate summary of disease dynamics for a pathogen such as HIV at the beginning of an outbreak (May et al., 2001). Some results can be inferred directly from the network topology but this is beyond the scope of this study.

In the case of PL, the basic reproduction number, denoted R_0 , increases with network size (Farkas et al., 2001). As R_0 and doubling time are negatively correlated in the early stages of an epidemic (May et al., 2001), doubling time is expected to decrease with population size (Newman, 2003; Keeling and Eames, 2005).

Further details are available in the Supplementary Methods of the Electronic Supplementary Material (ESM).

3. Results

We show that using biological assumptions to weight sexual contact networks strongly affects the epidemiology of STI. As the exact structure of these networks is the subject of an intense debate (Hamilton et al., 2008), we first present the results obtained on an empirical network (Potterat et al., 1985). However, our results hold when we consider theoretical finite size networks with PL and NB distributions.

3.1. Resulting realised number of sex acts

Our framework requires two biological assumptions to weight the network. The first assumption has to do with the allocation of the potential number of sex acts (one could think of this as the 'libido' of the individuals). The second has to do with the variance in the number of sex acts individuals have with each of their partners. For both these assumptions, existing data are very limited. However, one can evaluate the relevance of these assumptions by looking at the realised number of sex acts.

We find that for all allocations and partitioning, the total number of RSA and the total number of partners (degree) are

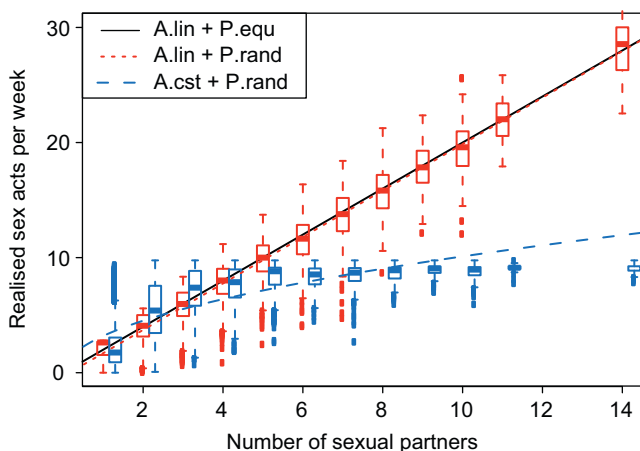


Fig. 2. Relationships between realised sex acts and number of partners for the empirical network. These curves show median values for 1000 stochastic SI simulations on the empirical network, where $\beta = 0.01$ and individuals have on average four sex acts per week. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

positively correlated, as shown in Fig. 2. While the assumptions used in most models (A_{lin} with either P_{max} or P_{rand}) yields a linear relationship between RSA and degree, the constant case (A_{cst} and P_{rand}) is characterised by a saturating slope for large number of partners. The increasing but saturating relationship in realised sex acts (RSA) as a function of degree observed in the case with constant allocation of sex acts is apparently counterintuitive. This paradox is explained by the fact that high degree nodes are less likely to loose potential sex acts (PSA) than low degrees nodes: even though the PSA are constant, a high-degree node is more likely to find partners to satisfy all its PSA whereas a low-degree node is less likely to find partners to use them. Note that the correction of the weighting by delta to keep the average number of RSA constant and equal to the average number of PSA is applied with the same intensity to each node of the corrected network: it is hence not changing the shape of the relationship between RSA (prior or after correction) and degree.

Empirical studies based on wide surveys tend to find that number of sex acts tend to increase with the number of partners, but that this increase saturates and is clearly not strictly proportional to the number of partners (Blower and Boe, 1993; Nordvik and Liljeros, 2006; Britton et al., 2007). This suggests that saturating or constant allocation models for PSA might be more realistic than the linear allocation, which is often implicitly assumed.

As explained in the Model section, any discrepancy in the total number of sex acts is corrected for in the following analyses. More precisely, we use δ to normalise the edge weights and thus keep the average number of RSA per individual constant for all the networks that we consider. Therefore, any change we observe in disease spread are only due to the weighted topology.

3.2. Disease spread on the empirical network

Fig. 3 shows the effect of biological assumptions on the topology of the network and how this affects disease spread. The 'classical' approach assumes the linear (A_{lin}) model for PSA allocation with equal (P_{equ}) or maximising (P_{max}) partitioning. As a consequence, all the interactions have the same weight and the number of RSA increases proportionally with the number of sexual partners (Fig. 3A). Assuming that the partitioning of the PSA among sexual partners is random (P_{rand}) generates variance in the weights of the edges (Fig. 3B). If we further assume that all the individuals are allocated the same number of potential sex acts (A_{cst}), we observe drastic changes. First, the variance in edge weights increases even more. Second, there is a complete shift in terms of sexual activities: nodes with few neighbours, whose activity was negligible with the classical approach, become important at the expenses of super-connected nodes (Fig. 3C). Fig. 3D illustrates the consequences of these assumptions on the spread of an STI (see the Methods for a description of the epidemiological model used). Assuming a P_{rand} partitioning model instead of a P_{equ} model slows the spread of the disease by more than 50% and assuming an A_{cst} allocation model instead of an A_{lin} model slows it by an extra 50%. Note that the longer the epidemics runs, the less realistic some of our model assumptions become (e.g. in the case of HIV, infected individuals are going to die).

To get a better understanding of these results, we analyse key variables on the same empirical network (Table 1). One of these variables is the median doubling time (t_d) of an STI. t_d estimates the speed of spread of a disease and we define it as the time required in stochastic simulations for the disease prevalence to go from 1% to 2% of the host population. We find that the median doubling time is significantly higher with a constant allocation model or with a random partitioning, when compared to the

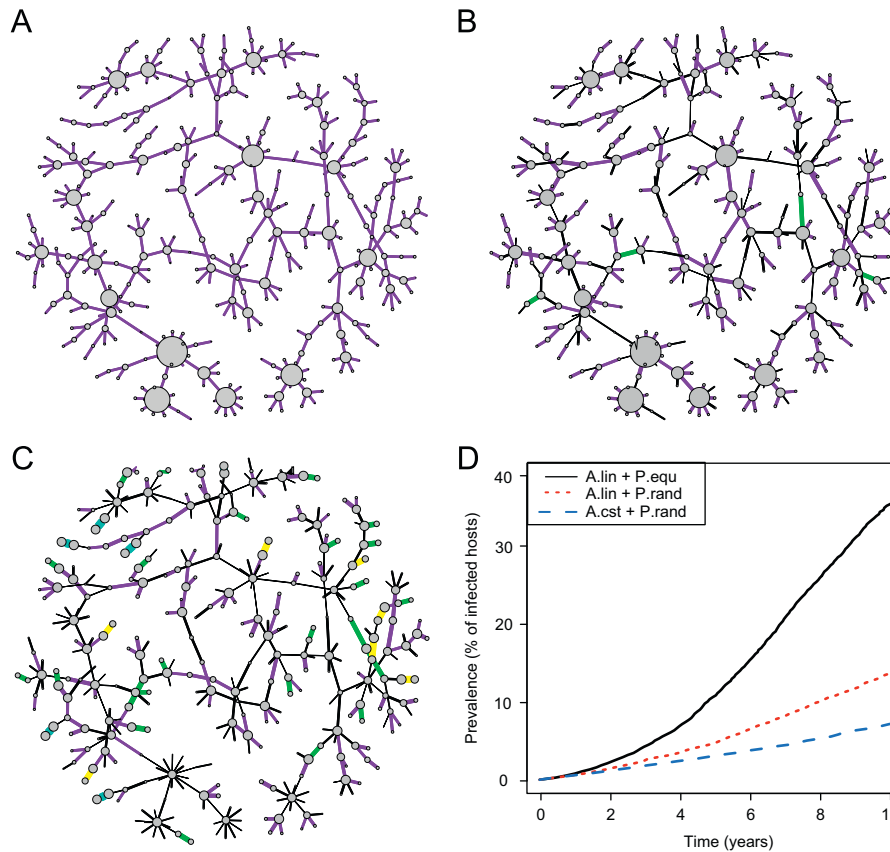


Fig. 3. Topological and epidemiological consequences of weighting the empirical network with realised sex acts. (A) Empirical contact network weighted with the number of realised sex acts implicitly assumed by classical models (A_{lin} and P_{equ}). (B) same as A but with a random partitioning, (C) same as A but with a constant allocation and a random partitioning. In panels A, B and C node diameter (edge width respectively) represents the number of realised sex acts per time step by an individual (the number of realised sex acts per time step by a couple respectively). Edge colour indicates less than 2 (in black), between 2 and 4 (purple), between 4 and 6 (green), between 6 and 8 (yellow) and more than 8 (cyan) sex acts per week. (D) Median STI prevalence as a function of time with the classical approach (in black), with a random partitioning (in red) and with a constant PSA allocation and a random partitioning (in blue). Notations and methods are as in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1

Effect of weighting the empirical network on disease spread. Each individual is allocated a number of PSA (lines) which can be partitioned in three different ways (columns). For each combination, the table shows the median doubling time (t_d , in weeks) with first and third quartile and δ , which is the ratio between the total number of realised sex acts (RSA) and the total number of potential sex acts (PSA) on the network. Edge weights are normalised by δ so that all the t_d are calculated on weighted networks with the same total number of RSA. We assume $\beta = 0.01$ and four RSA per week per individual on average. There were 1000 repetitions for each random partitioning. Further details are available in the Material and Methods and in the Supplementary Results.

Allocation	Estimator	Partitioning		
		P_{equ}	P_{max}	P_{rand}
A_{lin}	t_d	33 (19–57)	33 (19–57)	39 (22–79)
	δ	1.0	1.0	0.74 ± 0.01
A_{sat}	t_d	36 (23–61)	67 (37–130)	48 (28–92)
	δ	0.71	0.81	0.64 ± 0.01
A_{cst}	t_d	45 (27–76)	149 (72–404)	59 (34–109)
	δ	0.47	0.61	0.44 ± 0.01

classical approach (A_{lin} and P_{equ}). We also find an interaction between A_{cst} and P_{max} . Results obtained on 100 theoretical PL networks also show a significant effect of A_{sat} (see Supplementary Results).

As soon as we deviate from the classical approach (A_{lin} with P_{equ} or with P_{max}), the number of RSA over the whole network decreases. This is captured by δ , which is the ratio between the

total number of RSA and the total number of PSA over the whole network prior correction of the weighting. This decrease comes from the fact that super-connected individuals cannot match the availability of all their partners. Note that we use δ to normalise the edge weights and thus keep the average number of RSA per individual constant for all the networks that we consider.

3.3. Results on theoretical networks and HIV data

For the sake of generality, we tested our results on theoretical heterogeneous network of different sizes assuming two types of partner distributions: power-law (PL) and negative-binomial (NB). Further details about the theoretical networks are available in the Supplementary Methods. In Fig. 4, we show the results for 100 PL and 100 NB networks (see Supplementary results for a negative control using homogeneous small-world networks). In all cases, we find that diseases spread faster on networks weighted with the ‘classical’ A_{lin} PSA allocation model than on networks weighted using the A_{sat} model, which themselves lead to a faster spread than networks weighted using the A_{cst} model.

One of the consequences of the weighting is that the doubling times estimated at the beginning of an epidemic (Table 1) and the number of sex acts per week needed to reproduce early HIV prevalence observed throughout the world are better explained by weighted than unweighted PL networks (both on empirical or theoretical networks). The doubling times we obtain from the simulated evolution of epidemic prevalence on realistically weighted networks are in the same order of magnitude as that

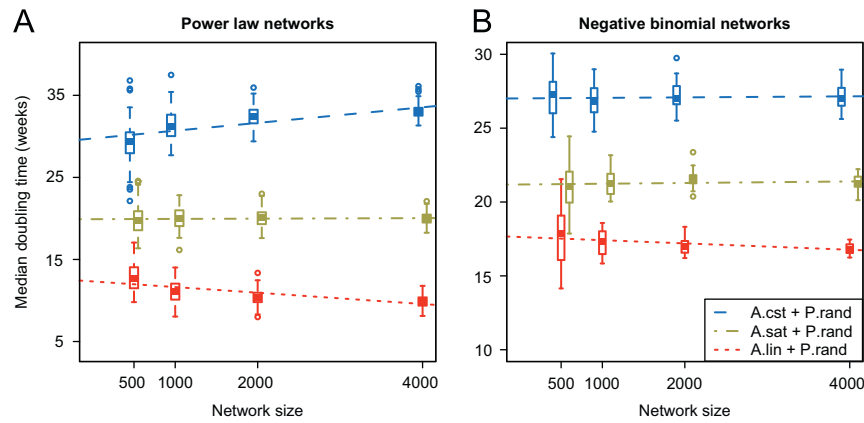


Fig. 4. Effect of weighting theoretical PL (A) and NB (B) networks of different sizes on disease spread. Disease spread is estimated through the median doubling time (t_d) computed over 100 runs of the simulation per replicate network. The box plot shows the median values over 100 different networks, the three quartiles and the outliers. Dashed lines are the output of a generalised linear model testing the effect of network size on t_d (see the main text). Parameters are as in Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

observed early in an HIV epidemic, i.e. approximately 1 year for men having sex with men (May and Anderson, 1987; Anderson and May, 1991). Note that the exact value of the doubling time is scaled by the product of the transmission rate per sex act (here set to $\beta = 0.01$; Vittinghoff et al., 1999; Boily et al., 2009) and the average number of realised sex acts per week per individual (here set to 4, see Supplementary Methods).

In the case of PL networks (Fig. 4A), we also observe that, with the A_{lin} PSA allocation model, disease doubling time decreases significantly with network size (in red, slope = $-6.9e-4$, p -value < 0.001), as expected for a network with a PL distribution (Farkas et al., 2001). For A_{sat} , however, the speed of disease spread is not affected by network size (in brown, slope = $3.2e-5$, p -value = 0.51). Finally, for A_{cst} , the spread is even slower on larger networks (in blue, slope = $9.5e-4$, p -value < 0.001). This result is interesting because, in the case of HIV, the doubling time of the disease has been shown to be constant whatever the geographic scale (city, country or continent) considered for the sexual network (May and Anderson, 1987, 1988; Anderson and May, 1991). The current explanation for this is that some sexual contact networks that seem to be PL are actually NB (Hamilton et al., 2008; Handcock and Jones, 2006). Our results suggest a new explanation, which is that even if the network node degree distribution follows a power law, edge weights can lead to the loss of decrease in doubling time with network size. Put differently, the mismatch between epidemiological data and the network structure could be due to the fact that using an unweighted network implicitly implies making unrealistic biological assumptions regarding the number of sex acts.

3.4. Effect of node removal on disease spread

Finally, we consider the effect on the median doubling time of removing a node on the empirical network (Fig. 5A). Node removal could be interpreted as any measure that blocks potential transmission. The difference between a prevention strategy and a treatment strategy is that node removal is independent from the host's infectious status. In other words, the goal is to modify the shape of the network in order to slow the speed of spread of an STI that may emerge in the population. Note that we are looking at doubling times so the higher the value, the more slowly the disease spreads. With the A_{lin} model, removing individuals with the highest number of partners (i.e. a targeted strategy) has the strongest effect on disease spread. This is captured by the slope between the degree of the node removed and the increase in doubling time: a zero slope means that there is no effect of targeting individuals with an increasing degree.

Here, we find a strongly increasing slope. This can be explained intuitively by the fact that super-connected individuals (or hubs) are also super-spreaders of the disease. For the A_{cst} model, targeting hubs has less impact than in the A_{lin} model (the slope of the regression between the change in median doubling time and the degree of the node removed drops from approximately 1/3 compared to the slope of the linear case). This suggests that other criteria might be more appropriate than node degree to detect key spreaders but none of the parameters we tested (number of sex acts realised, average degree of the neighbouring nodes, generalisation of this value to weighted networks; Eguiluz and Klemm, 2002; Barrat et al., 2004a, 2004b) performed better than the degree.

In a second procedure, we allow for multiple node removal (Fig. 5B and C). In this case, we compare the efficiency of a prevention policy aimed at individuals with many partners (in green) to a non-specific policy (in black). Again our goal is to see which modification of the network leads to a slower spread of a future emerging STI. We find that several individuals need to be removed for there to be a significant difference between these policies. If the PSA allocation is constant, prevention policies have less effect on the disease doubling time and more individuals need to be removed for there to be a difference between specific and non-specific targeting (see also Supplementary Results). This corroborates the results obtained in Fig. 4A by showing that in biologically realistic networks, aiming prevention policies at super-connected individuals might not yield a decrease in the spread of emerging STI as strong as one might expect. These results hold for theoretical networks (see Supplementary Results).

4. Discussion

Over the last decade, a passionate debate has emerged because sexual contact networks have been shown to have a highly heterogeneous structure (Liljeros et al., 2001; Hamilton et al., 2008; Schneeberger et al., 2004) but the spread of STI does not have the properties one would expect to find given such network structures (Hamilton et al., 2008; Handcock and Jones, 2006). Our results fit into this debate because they show that epidemiological trends in the data are consistent with the assumption that not all the edges of the network have the same weight. In other words, even though sexual networks do have a highly heterogeneous structure, they do not have associated properties because they are weighted and this weighting is not captured by studies that use contact-tracing or other methods to infer the shape of the network.

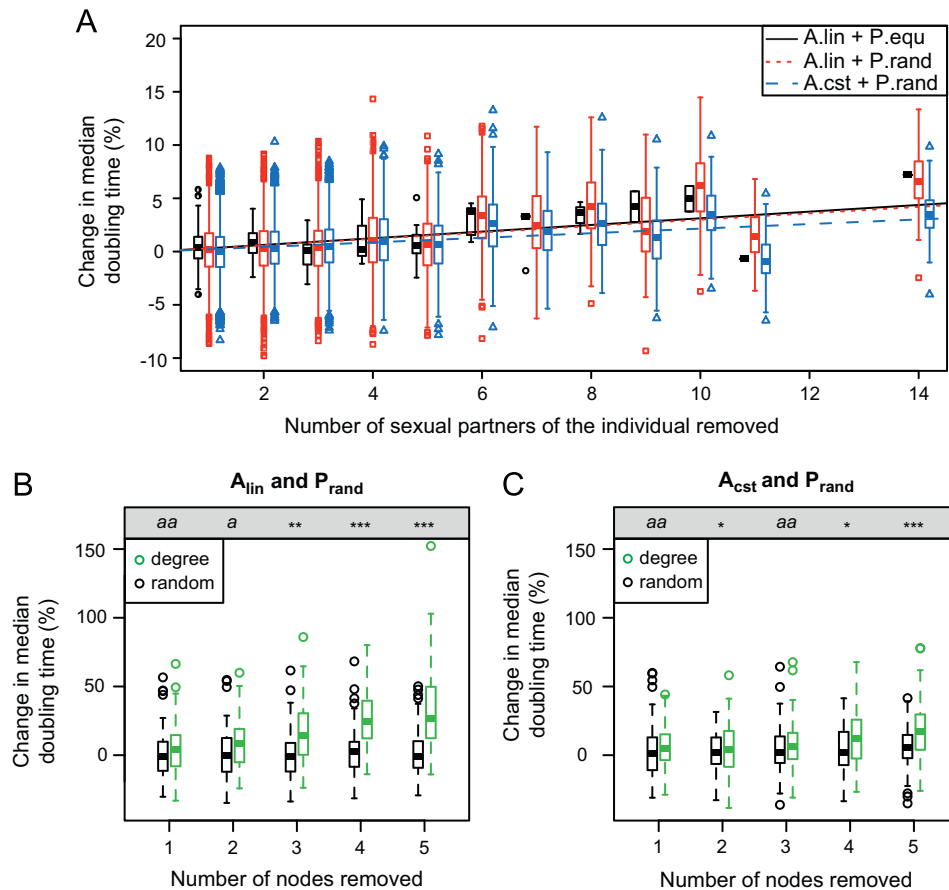


Fig. 5. Effect of prevention policies on disease spread on the empirical network. (A) Effect of single node removal on disease median doubling time as a function of node degree in the scenario A_{lin} with P_{equ} (the ‘classical’ case, in black), A_{lin} with P_{rand} (in red) or A_{cst} with P_{rand} (in blue). Dashed lines indicate the correlation found using a generalised linear model. Parameters are as in Fig. 2. Removing certain nodes increases disease spread (especially for a random partitioning) because we consider a prevention policy, which is done before the epidemic begins. (B) Effect of multiple node removal on disease median doubling time with the A_{lin} with P_{rand} scenario. In black, individuals are targeted at random (random) whereas in green the individuals with highest number of sexual partners (degree) are removed first. The grey box on the top indicates the p -value of a Wilcoxon signed-rank test comparing removal strategies. The significance code is ‘****’ for ≤ 0.001 , ‘***’ for ≤ 0.01 , ‘**’ for ≤ 0.05 and ‘a’ for non-significant but ≤ 0.1 and ‘aa’ for non-significant. (C) As panel B in the scenario with A_{cst} with P_{rand} . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We developed a framework to weight sexual contact networks using biological assumptions. This allows us to show that the underlying assumptions made when modelling disease spread on unweighted networks lead to correlation patterns between number of partners and number of sex acts that are at odds with empirical data (Blower and Boe, 1993; Nordvik and Liljeros, 2006). We also show that adopting a more realistic (weighted) network affects epidemiological dynamics in a way that is consistent with epidemiological data on HIV. Furthermore, these results hold for theoretical power-law networks and negative-binomial networks.

Our goal in this study is not to yield the most realistic model (we actually doubt that a decent level of realism can be achieved without adding a very large number of assumptions). Rather, our aim is to develop a general framework that is flexible enough to be applied to different epidemiological and sociological scenarios, while remaining biologically relevant. To the best of our knowledge, our allocation-partitioning model is first to present a biologically plausible relationship between total number of sex acts and total number of partners derived from explanatory behavioural hypotheses. In earlier studies (e.g., Newman, 2002; Joo and Lebowitz, 2004), this relationship is imposed without any mechanistic justification.

Interestingly the ratio of realised over potential number of sex acts (δ in Table 1) can yield insights into associated human behaviours (allocation and partitioning). This suggests that

combining our model with epidemiological data could allow us to investigate some sexual behaviours.

Optimal public health policies depend on the way diseases spread (Anderson and May, 1991). Even if weighted networks lose their heterogenous-related properties, removing the most connected nodes still has the strongest effect on the STI spread. However, if the prevention policy covers a too small fraction of the population, targeting super-spreader might not be a significantly better strategy than random targeting. This effect is accentuated by using a more realistic allocation model for sex acts. Of course, these results on the effect of public health should be taken with care, for instance because the epidemiological model we use is extremely simplified.

Other studies have addressed the problem of the discrepancy between network theory and HIV epidemiological data (see e.g. Röttingen and Garnett, 2002; Hamilton et al., 2008). In particular, a widespread method is to model ‘concurrent partnerships’ by considering that not all the edges of the network are active at the same time (Morris and Kretzschmar, 1997; Bauch and Rand, 2000; Eames and Keeling, 2004). Such models have increased biological realism but they require many additional assumptions that are difficult to test empirically. Moreover, they too make strong assumptions concerning the number of sex acts. In a future study, we will investigate the consequence of adding dynamical partnerships to a weighted network in order to disentangle the effect of the two on disease spread.

This study can be extended in other ways. First, other network topologies could be investigated (see also our Supplementary Results). We could consider more complicated epidemiological models and/or other STI. Also, it is very likely that people modify their behaviour once they know that they are infected with an STI (Funk et al., 2010). Finally, we focused on STI but our method can be applied to any type of networks with time-consuming interactions, e.g. in economics or in ecology (Newman, 2003).

Finally, this study calls for further studies on the link between the number of partners and the number of sex acts. Such data can be acquired without knowing the exact contact structure detail and it could greatly help to check whether the allocation and partitioning hypotheses of the model make sense. Furthermore, this study also underlines the importance of deriving new estimators to predict the speed of disease spread on weighted finite-size networks.

Acknowledgements

Many thanks to P. Bild, J. Cox, J. Heijne, M. Kenschake, R. Kouyos, H. Lentz, G. Leventhal, S. Lion, G. Paul, M. Pautasso, M. Salathé and T. Stadler for helpful discussions. We are very grateful to J.J. Potterat and S.Q. Muth for sharing their contact-network data. We thank the ETH Zürich, the ENS Cachan, the Île-de-France Regional Council under the MIDEM project in the framework DIM Malinf, and the French Ministries in charge of Agriculture and Environment for financial support. MML and SA performed the simulations and analysed the data, all three authors designed the study and wrote the paper. Electronic supplementary material (ESM) contains Supplementary Methods and Supplementary Results. SA is funded by an ATIP Avenir from CNRS and INSERM.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.jtbi.2012.06.031>.

References

- Althaus, C.L., Heijne, J.C.M., Roellin, A., Low, N., 2010. Transmission dynamics of *Chlamydia trachomatis* affect the impact of screening programmes. *Epidemics* 2 (3), 123–131.
- Anderson, R.M., May, R.M., 1991. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, New York, USA.
- Barabasi, A.L., Albert, R., 1999. Emergence of scaling in random networks. *Science* 286 (5439), 509–512. <http://dx.doi.org/10.1126/science.286.5439.509>.
- Barrat, A., Barthelemy, M., Vespignani, A., 2004a. Modeling the evolution of weighted networks. *Phys. Rev. E* 70 (6), 06149. <http://dx.doi.org/10.1103/PhysRevE.70.06149>.
- Barrat, A., Barthelemy, M., Pastor-Satorras, R., Vespignani, A., 2004b. The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. U.S.A.* 101 (11), 3747–3752. <http://dx.doi.org/10.1073/pnas.0400087101>.
- Bauch, C., Rand, D.A., 2000. A moment closure model for sexually transmitted disease transmission through a concurrent partnership network. *Proc. R. Soc. Lond. B* 267 (1456), 2019–2027. <http://dx.doi.org/10.1098/rspb.2000.1244>.
- Blower, S.M., Boe, C., 1993. Sex acts, sex partners, and sex budgets: implications for risk factor analysis and estimation of HIV transmission probabilities. *J. Acquir. Immune Defic. Syndr.* 6 (12), 1347–1352.
- Boily, M.C., Baggaley, R.F., Wang, L., Masse, B., White, R.G., Hayes, R.J., Alary, M., 2009. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. *Lancet Infect. Dis.* 9 (2), 118–129. [http://dx.doi.org/10.1016/S1473-3099\(09\)70021-0](http://dx.doi.org/10.1016/S1473-3099(09)70021-0).
- Britton, T., Nordvik, M.K., Liljeros, F., 2007. Modelling sexually transmitted infections: the effect of partnership activity and number of partners on R_0 . *Theor. Popul. Biol.* 72 (3), 389–399. <http://dx.doi.org/10.1016/j.tpb.2007.06.006>.
- Britton, T., Deijfen, M., Liljeros, F., 2011. A weighted configuration model and inhomogeneous epidemics. *J. Stat. Phys.* 145, 1368–1384. <http://dx.doi.org/10.1007/s10955-011-0343-3>.
- Catanzaro, M., Boguñá, M., Pastor-Satorras, R., 2005. Generation of uncorrelated random scale-free networks. *Phys. Rev. E* 71 (2), 027103. <http://dx.doi.org/10.1103/PhysRevE.71.027103>.
- Eames, K.T.D., Keeling, M.J., 2004. Monogamous networks and the spread of sexually transmitted diseases. *Math. Biosci.* 189, 115–130. <http://dx.doi.org/10.1016/j.mbs.2004.02.003>.
- Eames, K.T.D., Read, J.M., Edmunds, W.J., 2009. Epidemic prediction and control in weighted networks. *Epidemics* 1 (1), 70–76. <http://dx.doi.org/10.1016/j.epidem.2008.12.001>.
- Eguiluz, V.M., Klemm, K., 2002. Epidemic threshold in structured scale-free networks. *Phys. Rev. Lett.* 89 (10), 108701. <http://dx.doi.org/10.1103/PhysRevLett.89.108701>.
- Farkas, I.J., Derenyi, I., Barabasi, A.L., Vicsek, T., 2001. Spectra of 'real-world' graphs: beyond the semicircle law. *Phys. Rev. E* 6402 (2), 026704. <http://dx.doi.org/10.1103/PhysRevE.64.026704>.
- Ferguson, N.M., Garnett, G.P., 2000. More realistic models of sexually transmitted disease transmission dynamics: sexual partnership networks, pair models, and moment closure. *Sex. Transm. Dis.* 27 (10), 600–609.
- Foucault, M., 1976. *Histoire de la sexualité*. Gallimard, Paris, France.
- Funk, S., Salathé, M., Jansen, V.A.A., 2010. Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J. R. Soc. Interface* 7 (50), 1247–1256. <http://dx.doi.org/10.1098/rsif.2010.0142>.
- Hamilton, D.T., Handcock, M.S., Morris, M., 2008. Degree distributions in sexual networks: a framework for evaluating evidence. *Sex. Transm. Dis.* 35 (1), 30–40. <http://dx.doi.org/10.1097/OLQ.0b013e3181453a84>.
- Handcock, M.S., Jones, J.H., 2004. Likelihood-based inference for stochastic models of sexual network formation. *Theor. Popul. Biol.* 65 (4), 413–422. ISSN 0040-5809. <http://dx.doi.org/10.1016/j.tpb.2003.09.006>.
- Handcock, M.S., Jones, J.H., 2006. Interval estimates for epidemic thresholds in two-sex network models. *Theor. Popul. Biol.* 70 (2), 125–134. <http://dx.doi.org/10.1016/j.tpb.2006.02.004>.
- Joo, J., Lebowitz, J.L., 2004. Behavior of susceptible-infected-susceptible epidemics on heterogeneous networks with saturation. *Phys. Rev. E* 69 (6), 066105. <http://dx.doi.org/10.1103/PhysRevE.69.066105>.
- Keeling, M.J., Eames, K.T., 2005. Networks and epidemic models. *J. R. Soc. Interface* 2 (4), 295–307. <http://dx.doi.org/10.1098/rsif.2005.0051>.
- Liljeros, F., 2004. Sexual networks in contemporary Western societies. *Physica A* 338 (1–2), 238–245. ISSN 0378-4371. <http://dx.doi.org/10.1016/j.physa.2004.02.046>.
- Liljeros, F., Edling, C.R., Amaral, L.A.N., Stanley, H.E., Aberg, Y., 2001. The web of human sexual contacts. *Nature* 411 (6840), 907–908. <http://dx.doi.org/10.1038/35082140>.
- May, R.M., 2006. Network structure and the biology of populations. *Trends Ecol. Evol.* 21 (7), 394–399. <http://dx.doi.org/10.1016/j.tree.2006.03.013>.
- May, R.M., Anderson, R.M., 1987. Transmission dynamics of HIV infection. *Nature* 326 (6109), 137–142. <http://dx.doi.org/10.1038/326137a0>.
- May, R.M., Anderson, R.M., 1988. The transmission dynamics of human immunodeficiency virus (HIV). *Philos. Trans. R. Soc. B* 321 (1207), 565–607. <http://dx.doi.org/10.1098/rstb.1988.0108>.
- May, R.M., Gupta, S., McLean, A.R., 2001. Infectious disease dynamics: what characterizes a successful invader? *Philos. Trans. R. Soc. B* 356 (1410), 901–910.
- Morris, M., Kretzschmar, M., 1997. Concurrent partnerships and the spread of HIV. *AIDS* 11 (5), 641–648. ISSN 0269-9370.
- Newman, M.E.J., 2002. Spread of epidemic disease on networks. *Phys. Rev. E* 66 (1), 16128. <http://dx.doi.org/10.1103/PhysRevE.66.016128>.
- Newman, M.E.J., 2003. The structure and function of complex networks. *SIAM Rev.* 45 (2), 167–256.
- Nordvik, M.K., Liljeros, F., 2006. Number of sexual encounters involving intercourse and the transmission of sexually transmitted infections. *Sex. Transm. Dis.* 33 (6), 342–349. <http://dx.doi.org/10.1097/01.olq.0000194601.25488.b8>.
- Potterat, J.J., Rothenberg, R.B., Woodhouse, D.E., Muth, J.B., Pratts, C.I., Fogle, J.S., 1985. Gonorrhoea as a social disease. *Sex. Transm. Dis.* 12 (1), 25–32.
- Röttingen, J.-A., Garnett, G.P., 2002. The epidemiological and control implications of HIV transmission probabilities within partnerships. *Sex. Transm. Dis.* 29 (12), 818–827.
- Schneeberger, A., Mercer, C.H., Gregson, S.A.J., Ferguson, N.M., Nyamukapa, C.A., Anderson, R.M., Johnson, A.M., Garnett, G.P., 2004. Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. *Sex. Transm. Dis.* 31 (6), 380–387.
- Vittinghoff, E., Douglas, J., Judson, F., McKirnan, D., MacQueen, K., Buchbinder, S.P., 1999. Per-contact risk of human immunodeficiency virus transmission between male sexual partners. *Am. J. Epidemiol.* 150 (3), 306–311.
- Wang, H.Z., Liu, Z.R., Xu, H.H., 2007. Epidemic spreading on uncorrelated heterogeneous networks with non-uniform transmission. *Physica A* 382 (2), 715–721. <http://dx.doi.org/10.1016/j.physa.2007.04.034>.