Number of sex acts matters for heterosexual transmission and control of *Chlamydia trachomatis*

Christian L. Althaus[1,†], Marc Choisy[2,3], Samuel Alizon[2,†]
and the CSF group[*]

[1] Institute of Social and Preventive Medicine (ISPM),
University of Bern, 3012 Bern, Switzerland

[2] Laboratoire MIVEGEC (UMR CNRS 5290, IRD 224, UM1, UM2),
34394 Montpellier Cedex 5, France

[3] Oxford University Clinical Research Unit, Hanoi, Vietnam

[†] Both authors contributed equally to this work.

Author for correspondence: Christian L. Althaus (christian.althaus@alumni.ethz.ch)

Abstract

Mathematical models are instrumental in controlling *Chlamydia trachomatis* and other sexually transmitted infections (STIs). Modeling approaches that stratify the population by the number of individuals' sex partners often assume the transmission risk per partner to be constant. Sexual behavior data shows, however, that people with many partners share less sex acts per partner than people with fewer partners. This should lower the risk of transmission per partner for highly sexually active individuals and could have important epidemiological consequences for STI transmission and the projected impact of control scenarios. We devise a new epidemiological model that we fit to chlamydia prevalence data from Natsal-2 and CSF, two population-based probability sample surveys of sexual behavior in Britain and France. Compared to a standard model where the transmission risk per partner is constant, a model with realistic numbers of sex acts per partner provides a better fit to the data. Furthermore, the improved model provides evidence for strong assortative mixing among individuals with different numbers of sex partners. Our results suggest that all chlamydia infected individuals with one or more new heterosexual partners per year contribute to ongoing transmission, underlining that control interventions should be aimed towards all sexually active young adults.

Keywords: *Chlamydia trachomatis*, infection control, sexual mixing, mathematical model, parameter inference

1 Introduction

The spread of sexually transmitted infections (STIs) crucially depends on the number of sex partners and the number of sex acts within a sexual partnership. Empirical studies have shown that the number of sex partners within a given time period is a strong determinant for the spread of curable STIs such as *Chlamydia trachomatis* [1–3] and *Neisseria gonorrhoea* [4]. However, apart from a few noticeable exceptions, which are further discussed below, the number of sex acts as an additional factor for transmission has not received as much attention. In particular, mathematical models of STI transmission that are frequently used in guiding public health policy decision making often assume the number of sex acts per partner to be constant and independent on the number of an individual's partners [5–9]. Neglecting realistic numbers of sex acts per partner could have important implications for interpreting the results of STI transmission models, such as predicting the impact of screening interventions for chlamydia.

Theoretical studies have shown that taking into account the number of sex acts in models of STI transmission could be crucial [10, 11]. An important aspect was illustrated in the study by Nordvik & Liljeros [12], where they showed that it is not only the number of partners, but also the number of sex acts per partner, that is important for the transmission of an infection. Garnett & Anderson [13] developed a generic framework to include the heterogeneity in the number of sex acts between individuals and showed that it can enhance the likelihood of persistence of STIs. Still, we lack a solid understanding of the relationship between the number of sex acts and the number of sex partners. Furthermore, it remains to be determined how this relationship can be appropriately integrated into mathematical models of STI transmission that then can be validated by data.

The results of probability sample surveys of sexual behavior allow us to formulate some hypotheses on the relationship between the number of sex acts and the number of sex partners. Blower & Boe [14] analyzed data from 1770 unmarried individuals aged 20–44 years collected in San Francisco (USA) from 1988–1989. They concluded that the total number of sex act does not increase linearly with the number of sex partners and that the number of sex acts should be seen as a 'budget' split between sex partners. Nordvik & Liljeros [12] looked at the total number of sex acts in relation to the total number of sex partners based on data from 1150 individuals aged 16–31 years collected in Sweden in 1988. The authors found that the total number of sex acts remains constant, or in women even decreases, in individuals with higher numbers of sex partners. In principle, one can directly infer the relationship between sex acts and sex partners from sexual behavior surveys, but there remain potential biases due to the self-reported data. For example, it proves difficult to discriminate between protected and unprotected sex acts in such surveys. To date, we still lack a quantitative picture of the relationship between the number of sex partners and the number of sex acts, and its implications on the epidemiology of STIs.

Another important aspect of taking the number of sex acts between sex partners into account is its effect on the topology of the sexual contact network. If the number of sex acts per sex partner is constant, each contact (edge) between susceptible and infected individuals (nodes) is weighted equally, i.e., has the same transmission probability. This is not the case if sex acts are explicitly accounted for [15]. However, network-based models usually require computationally-intensive simulations, which is why mathematical models of STIs tend to adopt a simpler approach by stratifying the population into subgroups with varying levels of sexual activity [5, 7, 8]. These models are often referred

to as risk class models, though the stratification is based on the partner change rate that does not necessarily correlate with risk. The level of sexual mixing between individuals can vary between fully proportionate (to the number of sex partners individuals have), completely assortative (individuals within a sexual activity or risk group mix only with themselves) or somewhere in between. While some studies based on sexual behavior data indicate a certain level of assortative mixing between individuals [16–19], the output of STI transmission models is often only consistent with close to proportionate (or random) mixing [8, 20]. This inconsistency illustrates the challenge in estimating the degree of sexual mixing, and underlines the importance of considering the number of sex acts per sex partner in models of STI transmission.

Our goal is to gain a more detailed picture of the relationship between the number of heterosexual sex acts and the number of heterosexual sex partners and study its effect on the transmission of chlamydia, the most common bacterial STI. To this end, we make use of two data sets of sexual behavior: the British National Survey of Sexual Attitudes and Lifestyles (Natsal-2) [21], and the 'Contexte de la Sexualité en France' (CSF 2006) [22, 23]. First, we directly infer the relationship between the number of heterosexual sex acts and the number of heterosexual sex partners from Natsal-2. We then fit a novel epidemiological model of STI transmission to chlamydia prevalence data of both data sets and indirectly infer epidemiological parameters as well as the relationship between the number of heterosexual sex acts and heterosexual sex partners. Both methods support the notion that the number of heterosexual sex acts per partner decreases with higher number of heterosexual sex partners. This is also consistent with a high level of assortative mixing between individuals with similar number of sex partners. Lastly, we show the implications of taking realistic numbers of sex acts into account on modeling the heterosexual transmission dynamics of chlamydia and the impact of control interventions.

2 Materials and methods

2.1 Data

2.1.1 Natsal-2

Natsal-2 is a population-based probability sample survey of sexual attitudes and lifestyles conducted in Britain in 1999-2001 [21]. The sample consists of 11,161 women and men aged 16–44 years. Urine samples for ligase chain reaction (LCR) testing for chlamydia infection are available for a subset of 3569 sexually active respondents aged 18–44 years [1]. We use the following variables: number of new heterosexual sex partners in the last year, number of occasions of heterosexual sex in last 4 weeks, chlamydia test result from urine sample. Individual weights are used for all variables to adjust for unequal selection probabilities and to correct for the age and gender profile in the survey population.

2.1.2 CSF

'Contexte de la Sexualité en France' (CSF 2006) is a survey conducted in France among 12,364 randomly chosen individuals aged 18–69 years [22, 23]. Urine samples testing for chlamydia infection are available for a subset of 7407 sexually active respondents. The two variables we used are the number of new heterosexual sex partners in the last year and the chlamydia test result from urine sample. As for the Natsal-2 data, individual

weights are used to adjust for unequal selection probabilities and to correct for the age and gender profile in the population.

2.2 Sex acts and sex partners

We stratify the population by the reported number of new heterosexual partners in the last year which results in n different sexual activity classes with c_i sex partners. The average heterosexual activity of each individual during one year is assumed to be reflected in the reported number of heterosexual sex acts during the last 4 weeks. For simplicity, and in order to increase sample size, we pool data of women and men together.

The reported numbers of sex acts during the last 4 weeks are highly dispersed (figure 1A). Therefore, we assume that the observed number of sex acts during the last 4 weeks (Z) for a given individual with c_i new heterosexual partners during the last year is negative binomially distributed. The negative binomial distribution is frequently used to describe overdispersed data in biology and epidemiology [24, 25]. In the context of data on the number of sex acts, it can be interpreted as a process of sequential Bernoulli trials to reach sexual satisfaction with the parameters (k,p): for each of Z sex acts there is a probability p that it will not be sexually satisfying and individuals engage in new sex acts until satisfaction has been reached in k of them. The negative binomial distribution can also be interpreted as a mixture of Poisson distributions where the rates at which different individuals have sex are drawn from a gamma distribution with shape parameter k and scale parameter $p/(1-p)$. More generally, a single Poisson distribution is obtained for $k \to \infty$ and the negative binomial distribution is reduced to a geometric distribution for $k = 1$.

We assume that the mean of the negative binomial distribution is given by the following functional relationship:

$$m_i = d_1 + d_2 \ c_i^{d_3}, \tag{1}$$

where m_i and c_i are the average number of sex acts and the number of sex partners for individuals of sexual activity class i, respectively. m_i can then be scaled to obtain the expected number of heterosexual sex acts per year (M_i).

We construct a series of nested models by fixing some of the parameters to specific values. For example, if $d_2 = 0$, the total number of sex acts does not increase with higher numbers of sex partners. For $d_3 = 1$, the number of sex acts changes linearly with higher numbers of partners. The nested models then allow to perform hypothesis testing about different functional relationships between the number of sex acts and sex partners.

2.3 Epidemiological model

We assume the population to be stratified into n different sexual activity classes x_i with c_i new heterosexual partners per year [5, 7]. x_i denotes the proportion of individuals in sexual activity class i. On average, individuals change their sexual activity after one year, and are proportionally redistributed among all sexual activity classes. For simplicity, we assume that sexual activity and the natural history and transmission of the infection are the same in women and men. If y_i is the proportion of infected individuals in sexual activity class i, an SIS (susceptible-infected-susceptible) model (i.e., with no immunity)

can then be written as follows:

$$\frac{\mathrm{d}y_i}{\mathrm{d}t} = \mu \sum_{j=1}^{n} x_j y_j + (1 - y_i)c_i \sum_{j=1}^{n} b_{ij}\rho_{ij}y_j - (\gamma + \mu)y_i, \qquad (2)$$

where $1/\gamma$ and $1/\mu$ denote the average duration of infection and the average time spent in one sexual activity class, respectively. ρ_{ij} represents the elements of the mixing matrix that, following Garnett *et al.* [8], can be defined as

$$\rho_{ij} = \epsilon\delta_{ij} + (1 - \epsilon)\frac{c_j x_j}{\sum_{l=1}^{n} c_l x_l}, \qquad (3)$$

where δ_{ij} denotes the Kronecker delta (it is equal to 1 if $i = j$ and to 0 otherwise). Mixing can be varied between proportionate ($\epsilon = 0$) and fully assortative ($\epsilon = 1$).

The per partnership transmission probability b_{ij} can be considered as a Bernoulli process of repeated transmission probabilities per sex act:

$$b_{ij} = 1 - (1 - \beta)^{a_{ij}} \qquad (4)$$

where β denotes the transmission probability per sex act and a_{ij} the number of sex acts in a partnership between an individual of sexual activity class i and j. To construct the matrix a_{ij}, we use the following algorithm:

1. Define $s_i = \frac{M_i}{c_i}$ as the available number of sex acts per partner for an individual of sexual activity class i.

2. Set $j = n$, i.e., start with the highest sexual activity class.

3. Distribute s_j among all partners of sexual activity class $k = [1, j]$ proportional to the sexual mixing matrix ρ_{jk} and to what the partners are willing to share (s_k).

4. For $k = [1, j - 1]$, update s_k to only those available sex acts that have not been shared with individuals of sexual activity class j or higher.

5. If $j > 1$, set $j = j - 1$ and go to step 3. Otherwise, stop.

We calculate the basic reproduction number, R_0, for a given sexual activity class as follows: $\frac{c_i \sum_{j=1}^{n}(b_{ij}\rho_{ij})}{\gamma}$. R_0 for the entire population is calculated using the next generation method [26, 27]. An overview of all parameters and variables to describe the epidemiological model as well as the sexual behavior is given in table 1.

2.4 Maximum likelihood estimation

2.4.1 Direct method

The parameters d_1, d_2 and d_3 that describe the functional relationship between the number of sex acts and sex partners (Equation 1) and the dispersion parameter k from the negative binomial distribution are estimated using maximum likelihood and the function `mle2` from the package `bbmle` [28] for the R software environment for statistical computing [29].

2.4.2 Indirect method

We fit the epidemiological model to chlamydia prevalence data to obtain maximum likelihood estimates of behavioral and/or infection parameters [28]. Given a model-predicted prevalence y_i for sexual activity class i, the log-likelihood to find k_i positive tests in a sample of X_i individuals over all n sexual activity classes is

$$L(\beta, \gamma, \epsilon, d_1, d_2, d_3) = \sum_{i=1}^{n} \left(\log \binom{X_i}{k_i} + k_i \log y_i + (X_i - k_i) \log(1 - y_i) \right). \quad (5)$$

The model predicted prevalences for each sexual activity class, y_i, are obtained by running Equation 2 into steady-state using the function `ode` from the R package `deSolve` [30]. The parameter inference is performed using the function `mle2` from the package `bbmle` [28]. We use the minimization algorithm by Nelder & Mead [31] and the method SANN, which is a variant of simulated annealing [32]. The first method is very robust in finding local optima, whereas the latter usually performs better on rough likelihood surfaces but is computationally slow.

3 Results

3.1 Sex acts and sex partners (direct method)

The reported number of heterosexual sex acts during the last 4 weeks in Natsal-2 show an increasing trend for higher number of new heterosexual partners during the last year (figure 1A). Fitting a series of nested models to the data suggests that a *linear* relationship between the number of sex acts, m_i, and the number of sex partners, c_i, describes the data best (table 2 and figure 1B, solid line). The model where the number of sex acts for individuals with different numbers of sex partners is *constant* also fits the data well (figure 1B, dashed line). In contrast, the commonly used assumption where the total number of sex acts is strictly *proportional* to the number of sex partners is not consistent with the data (figure 1B, dotted line). We also fit an exponential relationship between the number of sex acts and sex partners (as suggested by Garnett & Anderson [13]) to the data. This results in a poorer fit (AIC = 17380.1) compared to the models in table 2.

3.2 Parameter inference (indirect method)

We perform two series of model fitting to indirectly infer parameters from chlamydia prevalence data. First, we only consider epidemiological parameters: the per sex act transmission probability β, the infectious duration $1/\gamma$ and the sexual mixing coefficient ϵ. Second, we focus on inferring the shape of the functional relationship between the number of sex acts and the number of sex partners (Equation 1) and estimate d_1, d_2, d_3 and ϵ. For the second approach, we assume that β and γ are fixed and informed by the literature. The rationale for choosing two separate approaches is that estimating all six parameters simultaneously would result in overfitting of the model.

3.2.1 Epidemiological parameters

We fit two models to the Natsal-2 and CSF chlamydia prevalence data to estimate the epidemiological parameters. In one model, the average number of sex acts for individuals

with a specific number of sex partners is given by the best fit model from the direct method (*Linear*). In the other model, we use the classical assumption where the number of sex acts increases strictly *proportional* with the number of sex partners. Both models fit the data well (figure 2) but vary in their estimates of the sexual mixing coefficient (table 3). While the *proportional* model suggests that sexual mixing is proportionate ($\epsilon = 0$), the *linear* model estimates the mixing coefficient to be more assortative and significantly different from zero ($\epsilon = 0.83$ (95% CI: 0.46–0.96) and $\epsilon = 0.53$ (95% CI: 0.38–0.68) for Natsal-2 and CSF, respectively). The estimated recovery rates and per sex act transmission probabilities for the *linear* model are in good agreement with previous estimates [33–36]. In contrast, the estimates of the per sex act transmission probability for the proportional model are very low, which highlights that the number of sex acts are not realistically taken into account.

3.2.2   Functional relationship

We infer the functional relationship of the number of sex acts and sex partners by fitting the model to the Natsal-2 and CSF chlamydia prevalence data simultaneously (table 4). To this end, we set the per sex act transmission probability and the infectious duration to values similar as estimated in table 3 ($\beta = 3\%$ and $1/\gamma = 1$ year). We find that the model where the number of sex acts does not increase with the number of partners (*Constant*) describes the data best. As in section 3.2.1, we again find evidence for assortative mixing ($\epsilon = 0.60$, 95% CI: 0.28–0.85). The estimates of d_2 or d_3 are close to zero for the model with *no intercept*, the *linear* model and the *full* model. This indicates that the number of sex acts does not substantially increase with higher numbers of partners. The *proportional* model clearly provides the worst fit to the data, and is the only model that suggests random sexual mixing. Interestingly, the number of estimated sex acts in the last four weeks is slightly higher than the estimate from the direct method (10.6 $vs.$ 7.0 for the constant model, respectively). This could either reflect an under-estimation of the per sex transmission probability or an under-reporting of sex acts in Natsal-2.

Overall, our indirect method to infer the functional relationship between the number of sex acts and sex partners corroborates the results from the direct method. Models where the number of sex acts is *constant* or show a slight *linear* increase with the number of sex partners provide the best and most parsimonious fit to the data. In contrast, the classical assumption, which assumes strict proportionality between the total number of sex acts and the number of sex partners is clearly worse in explaining the data.

3.3   Chlamydia transmission and control

We now investigate the effects of different assumptions about the number of sex acts between partners on chlamydia transmission and control. We focus on the general population in Britain (Natsal-2) and use the *linear* relationship for the number of sex acts as a function of sex partners from table 2. This 'realistic' model is contrasted to the classical assumption where the number of sex acts increases strictly *proportionally* with the number of sex partners. For both models, we use the best-fit estimates of the per sex act transmission probability β, the recovery rate γ and the sexual mixing coefficient ϵ from table 3.

For the *linear* model, distributing the number of sex acts to sex partners of different sexual activity classes results in the sex acts matrix a_{ij} (figure 3A). a_{ij} then allows to

calculate the per partnership transmission probabilities b_{ij} (equation 4) that range between 93.9% and 19.8% for partnerships between individuals with one and 10 partner(s), respectively (figure 3B). In contrast, the proportional model with a constant number of sex acts per partner results in a single per partnership transmission probability of 37.0%, which is close to what has been estimated for mid-risk populations [34].

Differences in the per partnership transmission probability will affect the transmission potential of individuals from different sexual activity classes. This is reflected in the basic reproduction number R_0 for each sexual activity class. In the classical scenario where the number of sex acts per partner is constant and the sexual mixing between partners is fully *proportional*, R_0 is a linearly increasing function of the number of new heterosexual partners per year (figure 4A, gray diamonds). The *proportional* model further suggests that only those individuals with three or more new heterosexual partners per year, or 21.2% of those individuals that have had a new partner in the previous year, will infect more than one additional person ($R_0 > 1$). In the more 'realistic' *linear* model, the value of R_0 exceeds the threshold of one for all individuals with one or more new heterosexual partners in the last year, and saturates around three for higher number of partners (figure 4A, black squares). The differences of the models is also reflected in the value of R_0 for the entire population. In the classical model, the R_0 for chlamydia is 1.26 while it is 1.12 for the more 'realistic' model.

Screening for and treating of asymptomatic chlamydia infection is the primary strategy to prevent disease transmission [37]. Ideally, screening will decrease R_0 below the threshold of one. The classical model predicts that low screening rates are sufficient to prevent transmission in individuals with low numbers of partners but that more than three tests per year on average would be necessary to prevent transmission in those individuals with many partners (right panel of figure 4, gray diamonds). Conversely, the 'more realistic' model illustrates that substantial screening (up to once every year) will already be necessary for individuals with low numbers of partners per year (right panel of figure 4B, black squares).

4   Discussion

This study provides insights into the relationship between the number of sex acts and the number of heterosexual partners, and how it affects the transmission dynamics of chlamydia and the projected impact of control interventions. We use direct and indirect methods to infer this relationship from two population-based probability sample surveys. Using the direct method, we find evidence that the number of sex acts increases only marginally with the number of sex partners. This finding is corroborated using the indirect method, where we show that a standard model in which the number of sex acts increases strictly proportionally with the number of sex partners, is less consistent with chlamydia prevalence data than most other models. Noticeably, the most 'realistic' model, where the number of sex acts are taken into account, provides strong evidence for assortative mixing between individuals with different numbers of heterosexual partners. The improved model also suggests that all individuals with one or more new heterosexual partners in the last year contribute to chlamydia transmission ($R_0 > 1$).

A major strength of this study is that we make use of two large data sets of sexual behavior. This allows us to perform a model comparison and reject those models that are not consistent with the data. We also use two methods to infer the relationship between the number of sex acts and the number of sex partners. The indirect method is based on chlamydia prevalence data and does not rely on self-reported number of sex

acts. Furthermore, this method only estimates potentially infectious, or unprotected, sex acts. Finally, we apply a novel epidemiological model of STI transmission that can take into the movement of individuals between different sexual activity classes.

There are several limitations to our study. First, we pool heterosexual women and men together, and do not take age-specific sexual behavior and different ethnic groups into account. Those simplifying assumptions are driven by the aim to keep the number of parameters small and the number of data points large enough to be able to statistically compare different models. This is also the reason why our analysis is focused on heterosexuals only. Data from larger probability sample surveys could potentially improve these limiting aspects. However, stratifying the population into women and men in the analysis using the direct method does not reveal significant differences between the two sexes (results not shown). Furthermore, we do not take concurrency of sexual partnerships, condom use and potential sex-specific differences in infection parameters into account.

We extend a commonly used STI transmission model, that stratifies the population into different sexual activity classes, with a sex acts matrix that describes the average numbers of sex acts between individuals of different sexual activity classes. The entries of our sex act matrix are either directly or indirectly inferred from data. Others have used similar models but could not derive the number of sex acts per partner in such detail [38, 39]. Somewhat different modeling frameworks that investigate the importance of sex acts by means of weighted contact networks have been developed by Britton *et al.* [40] and Moslonka-Lefebvre *et al.* [15]. These models are arguably more realistic than the more commonly used transmission models that stratify the population into different sexual activity classes. However, network models have two major disadvantages. First, the heterosexual contact networks for large populations are usually not known. Second, we currently lack meaningful summary statistics for weighted networks, and model results tend to rely on computationally expensive simulations. An alternative extension of our analyses would be to use a recently developed framework, that manages to derive analytical results for weighted networks by adopting a configuration network approach and using joint probability distributions of number of sex partners and number of sex acts [41].

Our study provides evidence for strong assortative mixing between heterosexual individuals of different sexual activity classes in the general population. Using the *linear* relationship between the number of sex acts and sex partners from the direct method, we estimate a sexual mixing coefficient of $\epsilon = 0.83$ for Natsal-2 (table 3). An early study based on partner notification data indicated that a proportional mixing model is not an adequate description of the observed sexual mixing pattern [42]. Later, Renton *et al.* [16] indicated that individuals with high rates of sexual partners preferentially select other individuals with high rates as partners. Garnett *et al.* [18] also found that sexual mixing was weakly assortative in patients attending sexually transmitted diseases (STDs) clinics. In contrast, Stoner *et al.* [43] did not find evidence for assortative mixing in members of gonorrhoea and chlamydia networks. This illustrates the difficulties in quantifying the degree of sexual mixing, and that the estimates are likely to depend on the studied population.

Our estimates of the basic reproduction number R_0 for chlamydia are consistent with values that were investigated in a recent theoretical study that introduced the concepts of case and partnership reproduction numbers [44]. Another study based on contact tracing data showed that individual reproduction numbers for chlamydia rarely exceed 3.0 [45]. Interestingly, this is exactly what we find for the host-specific basic reproduction

numbers from the *linear* model (figure 4A, black squares) where R_0 saturates around 3.0 for those individuals with a high number of sex partners.

We can show that taking into account the number of sex acts between partners with different levels of sexual activity has important consequences for modeling the transmission of STIs. First, we find strong evidence for assortative mixing between individuals of different heterosexual activity classes. Second, our results suggest that all chlamydia infected individuals with one or more new heterosexual partners per year contribute to ongoing transmission. Both of those findings have important implications for guiding public health decisions about chlamydia screening programs. A high level of assortative mixing means that chlamydia can easily persist in those sub-populations that are difficult to reach through screening. But our results also show that chlamydia control interventions should be aimed towards all sexually active young adults as they are likely to contribute to ongoing transmission.

5 Competing interests

We declare no competing interests.

6 Authors' contributions

CLA and SA designed the study, carried out the analysis, interpreted the results and drafted the manuscript. MC contributed to the analysis. All authors contributed to revising the manuscript and gave final approval for publication.

7 Acknowledgments

8 Funding

References

1. Fenton, K. A., Korovessis, C., Johnson, A. M., McCadden, A., McManus, S., Wellings, K., Mercer, C. H., Carder, C., Copas, A. J. *et al.* 2001 Sexual behaviour in Britain: reported sexually transmitted infections and prevalent genital Chlamydia trachomatis infection. *Lancet* **358**, 1851–1854.

2. Sonnenberg, P., Clifton, S., Beddows, S., Field, N., Soldan, K., Tanton, C., Mercer, C. H., da Silva, F. C., Alexander, S. *et al.* 2013 Prevalence, risk factors, and uptake of interventions for sexually transmitted infections in Britain: findings from the National Surveys of Sexual Attitudes and Lifestyles (Natsal). *Lancet* **382**, 1795–806. doi:10.1016/S0140-6736(13)61947-9.

3. Althaus, C. L., Turner, K. M. E., Schmid, B. V., Heijne, J. C. M., Kretzschmar, M. & Low, N. 2012 Transmission of Chlamydia trachomatis through sexual partnerships: a comparison between three individual-based models and empirical data. *J R Soc Interface* **9**, 136–146.

4. Ison, C. A., Town, K., Obi, C., Chisholm, S., Hughes, G., Livermore, D. M., Lowndes, C. M. & GRASP collaborative group 2013 Decreased susceptibility to cephalosporins among gonococci: data from the Gonococcal Resistance to Antimicrobials Surveillance Programme (GRASP) in England and Wales, 2007-2011. *Lancet Infect Dis* **13**, 762–8. doi:10.1016/S1473-3099(13)70143-9.

5. Hethcote, H. W. & Yorke, J. A. 1984 *Gonorrhea transmission dynamics and control.* Lecture Notes in Biomathematics. Berlin: Springer-Verlag.

6. Anderson, R. M. & May, R. M. 1991 *Infectious Diseases of Humans: Dynamics and Control.* Oxford: Oxford University Press.

7. Garnett, G. P. & Anderson, R. M. 1993 Factors controlling the spread of HIV in heterosexual communities in developing countries: patterns of mixing between different age and sexual activity classes. *Philos Trans R Soc Lond B Biol Sci* **342**, 137–59.

8. Garnett, G. P., Mertz, K. J., Finelli, L., Levine, W. C. & St Louis, M. E. 1999 The transmission dynamics of gonorrhoea: modelling the reported behaviour of infected patients from Newark, New Jersey. *Philos Trans R Soc Lond B Biol Sci* **354**, 787–97.

9. Keeling, M. J. & Rohani, P. 2008 *Modeling Infectious Diseases in Humans and Animals.* Princeton: Princeton University Press.

10. Kaplan, E. H. 1990 Modeling HIV infectivity: must sex acts be counted? *J Acquir Immune Defic Syndr* **3**, 55–61.

11. Røttingen, J.-A. & Garnett, G. P. 2002 The epidemiological and control implications of HIV transmission probabilities within partnerships. *Sex Transm Dis* **29**, 818–27.

12. Nordvik, M. K. & Liljeros, F. 2006 Number of sexual encounters involving intercourse and the transmission of sexually transmitted infections. *Sex Transm Dis* **33**, 342–349.

13. Garnett, G. P. & Anderson, R. M. 1996 Sexually transmitted diseases and sexual behavior: insights from mathematical models. *J Infect Dis* **174 Suppl 2**, S150–61.

14. 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**, 1347–52.

15. Moslonka-Lefebvre, M., Bonhoeffer, S. & Alizon, S. 2012 Weighting for sex acts to understand the spread of STI on networks. *J Theor Biol* **311**, 46–53. doi:10.1016/j.jtbi.2012.06.031.

16. Renton, A., Whitaker, L., Ison, C., Wadsworth, J. & Harris, J. R. 1995 Estimating the sexual mixing patterns in the general population from those in people acquiring gonorrhoea infection: theoretical foundation and empirical findings. *J Epidemiol Community Health* **49**, 205–13.

17. Garnett, G. P. & Anderson, R. M. 1993 Contact tracing and the estimation of sexual mixing patterns: the epidemiology of gonococcal infections. *Sex Transm Dis* **20**, 181–91.

18. Garnett, G. P., Hughes, J. P., Anderson, R. M., Stoner, B. P., Aral, S. O., Whittington, W. L., Handsfield, H. H. & Holmes, K. K. 1996 Sexual mixing patterns of patients attending sexually transmitted diseases clinics. *Sex Transm Dis* **23**, 248–57.

19. Aral, S. O., Hughes, J. P., Stoner, B., Whittington, W., Handsfield, H. H., Anderson, R. M. & Holmes, K. K. 1999 Sexual mixing patterns in the spread of gonococcal and chlamydial infections. *Am J Public Health* **89**, 825–33.

20. Althaus, C. L., Heijne, J. C., Herzog, S. A., Roellin, A. & Low, N. 2012 Individual and population level effects of partner notification for Chlamydia trachomatis. *PLoS One* **7**, e51 438.

21. Johnson, A. M., Mercer, C. H., Erens, B., Copas, A. J., McManus, S., Wellings, K., Fenton, K. A., Korovessis, C., Macdowall, W. *et al.* 2001 Sexual behaviour in Britain: partnerships, practices, and HIV risk behaviours. *Lancet* **358**, 1835–1842.

22. Bajos, N. & Bozon, M. 2006 Enquête Contexte de la Sexualité en France, 2006. Tech. rep., INSERM-INED.

23. Bajos, N. & Bozon, M. (eds.) 2008 *Enquête sur la sexualité en France: pratiques, genre et santé.* La Découverte, Paris.

24. Lloyd-Smith, J. O. 2007 Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. *PLoS One* **2**, e180.

25. Hamilton, D. T., Handcock, M. S. & Morris, M. 2008 Degree distributions in sexual networks: a framework for evaluating evidence. *Sex Transm Dis* **35**, 30–40.

26. Diekmann, O., Heesterbeek, J. A. & Metz, J. A. 1990 On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *J Math Biol* **28**, 365–82.

27. Diekmann, O. & Heesterbeek, J. A. P. 2000 *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis, and Interpretation.* Chichester: John Wiley.

28. Bolker, B. M. 2008 *Ecological models and data in R.* Princeton, NJ: Princeton University Press.

29. R Development Core Team 2009 *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

30. Soetaert, K., Petzoldt, T. & Setzer, R. W. 2010 Solving differential equations in R: package deSolve. *Journal of Statistical Software* **33**, 1–25.

31. Nelder, J. A. & Mead, R. 1965 A simplex method for function minimization. *The Computer Journal* **7**, 308–313. doi:10.1093/comjnl/7.4.308.

32. Bélisle, C. J. P. 1992 Convergence Theorems for a Class of Simulated Annealing Algorithms on Rd. *Journal of Applied Probability* **29**, pp. 885–895.

33. Althaus, C. L., Heijne, J. C., Roellin, A. & Low, N. 2010 Transmission dynamics of Chlamydia trachomatis affect the impact of screening programmes. *Epidemics* **2**, 123 – 131.

34. Althaus, C. L., Heijne, J. C. M. & Low, N. 2012 Towards More Robust Estimates of the Transmissibility of Chlamydia trachomatis. *Sex Transm Dis* **39**, 402–4.

35. Price, M. J., Ades, A. E., Angelis, D. D., Welton, N. J., Macleod, J., Soldan, K., Turner, K., Simms, I. & Horner, P. J. 2013 Mixture-of-exponentials models to explain heterogeneity in studies of the duration of Chlamydia trachomatis infection. *Stat Med* **32**, 1547–60. doi:10.1002/sim.5603.

36. Davies, B., Anderson, S.-J., Turner, K. M. E. & Ward, H. 2014 How robust are the natural history parameters used in chlamydia transmission dynamic models? A systematic review. *Theor Biol Med Model* **11**, 8. doi:10.1186/1742-4682-11-8.

37. Althaus, C., Turner, K., Mercer, C., Auguste, P., Roberts, T., Bell, G., Herzog, S., Cassell, J., Edmunds, J. *et al.* 2014 Effectiveness and cost-effectiveness of traditional and new partner notification technologies for curable sexually transmitted infections: observational study, systematic reviews and mathematical modelling. *Health Technol Assess* **18**. doi:10.3310/hta18020.

38. Regan, D. G., Wilson, D. P. & Hocking, J. S. 2008 Coverage is the key for effective screening of Chlamydia trachomatis in Australia. *J Infect Dis* **198**, 349–358.

39. Garnett, G. P., White, P. J. & Ward, H. 2008 Fewer partners or more condoms? Modelling the effectiveness of STI prevention interventions. *Sex Transm Infect* **84 Suppl 2**, ii4–11.

40. Britton, T., Nordvik, M. K. & Liljeros, F. 2007 Modelling sexually transmitted infections: the effect of partnership activity and number of partners on R0. *Theor Popul Biol* **72**, 389–99.

41. Kamp, C., Moslonka-Lefebvre, M. & Alizon, S. 2013 Epidemic spread on weighted networks. *PLoS Comput Biol* **9**, e1003 352. doi:10.1371/journal.pcbi.1003352.

42. Granath, F., Giesecke, J., Scalia-Tomba, G., Ramstedt, K. & Forssman, L. 1991 Estimation of a preference matrix for women's choice of male sexual partner according to rate of partner change, using partner notification data. *Math Biosci* **107**, 341–8.

43. Stoner, B. P., Whittington, W. L., Hughes, J. P., Aral, S. O. & Holmes, K. K. 2000 Comparative epidemiology of heterosexual gonococcal and chlamydial networks: implications for transmission patterns. *Sex Transm Dis* **27**, 215–23.

44. Heijne, J. C. M., Herzog, S. A., Althaus, C. L., Low, N. & Kretzschmar, M. 2013 Case and partnership reproduction numbers for a curable sexually transmitted infection. *J Theor Biol* **331**, 38–47. doi:10.1016/j.jtbi.2013.04.010.

45. Potterat, J. J., Zimmerman-Rogers, H., Muth, S. Q., Rothenberg, R. B., Green, D. L., Taylor, J. E., Bonney, M. S. & White, H. A. 1999 Chlamydia transmission: concurrency, reproduction number, and the epidemic trajectory. *Am J Epidemiol* **150**, 1331–9.

Table 1: Parameters and variables used to describe the sexual behavior and the epidemiological model.

Notation	Description
n	Number of host classes with different sexual activity
c_i	Number of new heterosexual sex partners in the last year for individuals of host class i
m_i	Average number of heterosexual sex in the last 4 weeks for individuals of host class i
M_i	Average number of heterosexual sex in the last year for individuals of host class i
d_1, d_2, d_3	Parameters describing the functional relationship between c_i and m_i
x_i	Proportion of the host population in sexual activity class i
X_i	Number of individuals in sexual activity class i
y_i	Proportion of infected individuals in sexual activity class i
μ	Rate at which individuals change their sexual activity
γ	Recovery rate of the infection
ρ_{ij}	Mixing matrix between individuals of sexual activity classes i and j
ϵ	Sexual mixing coefficient (assortative index)
δ_{ij}	Kronecker delta (a function equal to 1 if $i = j$, and to 0 otherwise)
s_i	Average number of sex acts per sex partner for an individual of host class i ($\frac{M_i}{c_i}$)
a_{ij}	Number of sex acts in a partnership between individuals of sexual activity classes i and j
β	Transmission probability per sex act
b_{ij}	Transmission probability per partnership between individuals of sexual activity classes i and j

Table 2: Maximum likelihood estimates of the functional relationship between the number of sex acts and the number of sex partners (direct method). Fixed parameters are given in squared brackets and estimates are shown together with 95% confidence intervals. The six different models were sorted in increasing order of their AIC. Only models whose AIC were within 10 from the best fit (*Linear*) model are shown together with the classical model where the number of sex acts increases strictly *proportional* to the number of sex partners. Based on likelihood ratio tests, the linear model is superior to the constant and full model ($p < 0.05$).

| Model | Parameters | | | | No. of free parameters | -2LogLik | ΔAIC |
	d_1	d_2	d_3	k			
Linear	6.64 (6.20, 7.09)	0.17 (0.01, 0.33)	[1.00]	0.65 (0.61, 0.69)	3	16821.9	0.0
Full	6.77 (6.29, 7.25)	0.07 (-0.14, 0.28)	1.34 (0.31, 2.36)	0.65 (0.61, 0.69)	4	16821.4	1.6
Exponent	5.77 (5.40, 6.15)	[1.00]	0.39 (0.13, 0.65)	0.65 (0.61, 0.69)	3	16824.3	2.4
No intercept	[0.00]	6.81 (6.42, 7.20)	0.06 (-0.01, 0.13)	0.65 (0.61, 0.69)	3	16825.6	3.7
Constant	7.01 (6.67, 7.34)	[0.00]	–	0.65 (0.61, 0.69)	2	16828.4	4.6
Proportional	[0.00]	5.08 (4.81, 5.35)	[1.00]	0.54 (0.51, 0.57)	2	17296.0	472.2

Table 3: Maximum likelihood estimates of epidemiological parameters. The relationship between the number of sex acts and the number of sex partners is either based on the best fit model from the direct method (*Linear*) or the classical assumption where the number of sex acts increases strictly *Proportional* to the number of sex partners. Estimates are shown together with 95% confidence intervals.

Model	Data	Parameters β	$1/\gamma$	ϵ	No. of free parameters	-2LogLik	AIC
Linear	Natsal-2	3.0% (1.9%, 4.1%)	1.12 y (0.83, 1.42)	0.83 (0.46, 0.96)	3	39.3	45.3
Linear	CSF	3.4% (NA, NA)	0.96 y (NA, NA)	0.53 (0.38, 0.68)	3	85.2	91.2
Proportional	Natsal-2	0.7% (0.3%, 1.1%)	1.16 y (0.44, 1.88)	0.00 (0.00, 0.00)	3	41.0	47.0
Proportional	CSF	0.4% (0.2%, 0.5%)	1.04 y (0.45, 1.63)	0.00 (0.00, 0.00)	3	105.7	111.7

Table 4: Maximum likelihood estimates of the functional relationship between the number of sex acts and sex partners (indirect method). Fixed parameters are given in squared brackets and estimates are shown together with 95% confidence intervals. The five different models are sorted in increasing order of their AIC. The models are fitted with initial parameter values from Table 2 and $\epsilon = 0.5$.

| Model	Parameters				No. of free	-2LogLik	AIC
	d_1	d_2	d_3	ϵ	parameters		
Constant	10.6 (8.54, 12.9)	[0.0]	[1.0]	0.60 (0.28, 0.85)	2	125.4	129.4
No Intercept	[0.0]	10.4 (6.44, 15.2)	5.1e-4 (0, 1.46)	0.63 (0.16, 0.94)	3	125.4	131.5
Linear	10.0 (7.92, 12.4)	4.6e-3 (0, 0.15)	[1.0]	0.69 (0.34, 0.89)	3	125.8	131.8
Full	6.72 (2.54, 12.8)	3.97 (0.59, 10.2)	1.1e-4 (0, 0.17)	0.59 (0.31, 0.82)	4	125.4	133.4
Proportional	[0.0]	1.28 (1.23, 1.34)	[1.0]	7.6e-3 (9.8e-5, 0.38)	2	153.2	157.2
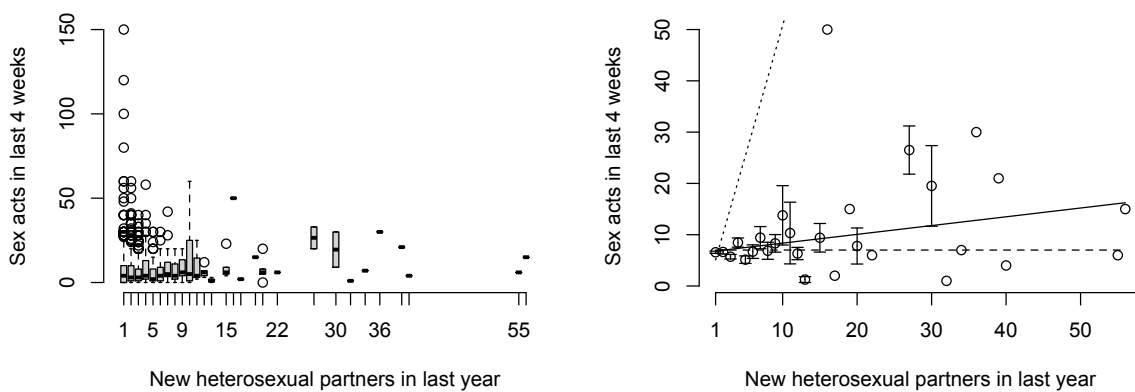
Figure 1: Number of heterosexual sex acts during the last 4 weeks as a function of the number of new heterosexual partners during the last year. (A) Boxplot of the reported data in Natsal-2 ($n = 2824$). The number of sex acts are positively correlated with the number of sex partners (Pearson's $r = 0.07$; $p < 10^{-3}$). (B) Functional relationship between the number of sex acts and the number of sex partners together with the means and standard errors of the reported data. The best fit model (*Linear*, solid line) is shown together with the models assuming the total number of sex acts (*Constant*, dashed line) or the number of sex acts per partner (*Proportional*, dotted line) to be constant.
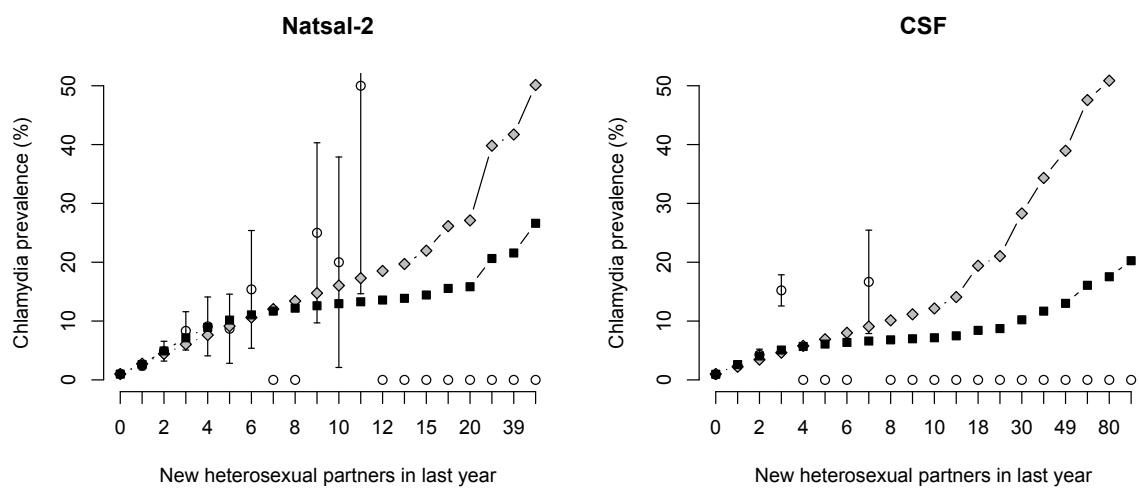
Figure 2: Chlamydia prevalence and model fits. For both data sets, the model where the total number of sex acts increases *linearly* with the number of sex partners (black squares) provides a better fit to the data compared to the classical model where the number of sex acts increases strictly *proportional* with the number of sex partners (gray diamonds).
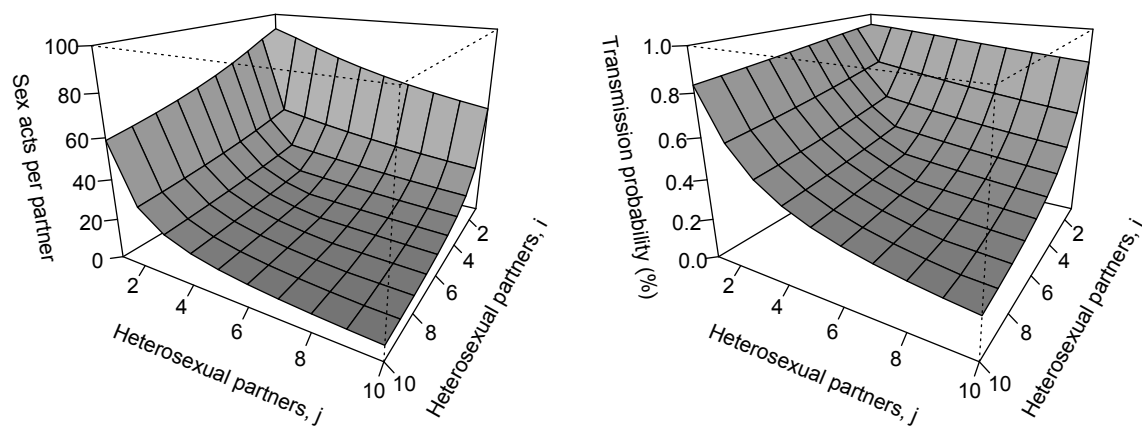
Figure 3: Average number of sex acts (left panel) and per partnership transmission probabilities (right panel) between individuals of sexual activity class i and j. The matrices correspond to the best-fit (*Linear*) model from section 3.2.1.
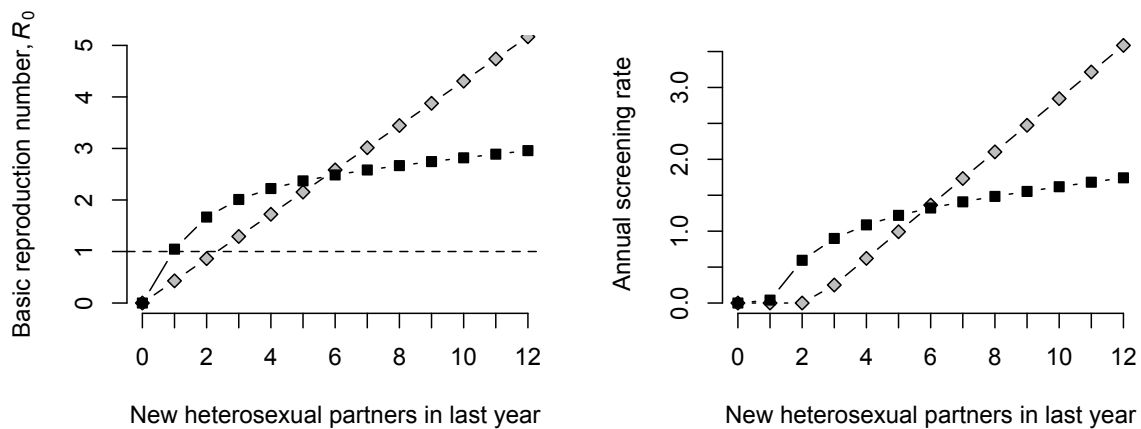
Figure 4: Host-specific basic reproduction number (left panel) and required screening rate to push R_0 below one (right panel). The results from the best-fit *linear* model (black squares) are compared to the classical model where the number of sex acts increases strictly *proportional* with the number of sex partners (gray diamonds). Data are from Natsal-2 only.