



## Age difference between heterosexual partners in Britain: Implications for the spread of *Chlamydia trachomatis*

Joost H. Smid<sup>a,\*</sup>, Victor Garcia<sup>a,1</sup>, Nicola Low<sup>a</sup>, Catherine H. Mercer<sup>b</sup>, Christian L. Althaus<sup>a</sup>

<sup>a</sup> University of Bern, Institute of Social and Preventive Medicine (ISPM), Switzerland

<sup>b</sup> UCL Research Department of Infection and Population Health, London, UK

### ARTICLE INFO

#### Keywords:

Chlamydia trachomatis  
Sexually transmitted diseases  
Age disparity  
Sexual behaviour  
Mathematical model

### ABSTRACT

Heterosexual partners often differ in age. Integrating realistic patterns of sexual mixing by age into dynamic transmission models has been challenging. The effects of these patterns on the transmission of sexually transmitted infections (STI) including *Chlamydia trachomatis* (chlamydia), the most common bacterial STI are not well understood. We describe age mixing between new heterosexual partners using age- and sex-specific data about sexual behavior reported by people aged 16–63 years in the 2000 and 2010 British National Surveys of Sexual Attitudes and Lifestyles. We incorporate mixing patterns into a compartmental transmission model fitted to age- and sex-specific, chlamydia positivity from the same surveys, to investigate *C. trachomatis* transmission. We show that distributions of ages of new sex partners reported by women and by men in Britain are not consistent with each other. After balancing these distributions, new heterosexual partnerships tend to involve men who are older than women (median age difference 2, IQR –1, 5 years). We identified the most likely age combinations of heterosexual partners where incident *C. trachomatis* infections are generated. The model results show that in > 50% of chlamydia transmitting partnerships, at least one partner is ≥25 years old. This study illustrates how sexual behavior data can be used to reconstruct detailed sexual mixing patterns by age, and how these patterns can be integrated into dynamic transmission models. The proposed framework can be extended to study the effects of age-dependent transmission on incidence in any STI.

### 1. Introduction

Heterosexual partners often differ in age. Men tend to have younger sexual partners (Ford et al., 2001; Kraut-Becher and Aral, 2006) but the magnitude and structure of these age differences between heterosexual partners are themselves age dependent (Gregson et al., 2002; Mercer et al., 2009) and vary between countries (Bajos et al., 1995; Beauclair et al., 2016; Gregson et al., 2002). How these age disparities affect the risk of acquiring and transmitting sexually transmitted infections (STI) remains poorly understood and evidence about their influence is mixed (Anderson et al., 1992; Garnett and Anderson, 1996). For example, the frequency of partnerships between younger women and older men in sub-Saharan Africa has been suggested to contribute to the high HIV infection prevalence in young women (Beauclair et al., 2016; Gregson et al., 2002; Leclerc-Madlala, 2008; UNAIDS, 2016). A review of studies of other STI in women found strong associations between older partner age and STI prevalence but not incidence (Swartzendruber et al., 2013). Detailed descriptions of age mixing patterns combined with pathogen-specific epidemiologic information could improve understanding of the

risk of STI in a population (Kraut-Becher and Aral, 2006; Ott et al., 2011).

Descriptions of age mixing patterns depend on both the data collection approach and how those data are analyzed. Cross-sectional surveys often show discrepancies between what men and women report about their partner's age, which might be influenced by social desirability biases, reflecting societal expectations about sexual partnerships (Fenton et al., 2001). A study in which both partners in a couple were interviewed, in Malawi, found that women tended to underestimate their partner's age (with inaccuracy increasing with age disparity) (Helleringer et al., 2011). The variability of age difference discrepancies between heterosexual partners suggests that summarizing these disparities by averaging offers an incomplete description of age mixing. Information about the variability in age differences, which can be considerable and increases with increasing age, is also needed (Mercer et al., 2009).

Mathematical models offer a useful approach to investigate how age mixing patterns affect the incidence, prevalence and spread of STI, but modeling realistic patterns of sexual mixing by age has been

\* Corresponding author at: Finkenhubelweg 11, 3012, Bern, Switzerland.

E-mail address: [joost.smid@ispm.unibe.ch](mailto:joost.smid@ispm.unibe.ch) (J.H. Smid).

<sup>1</sup> Both authors have equally contributed to this work.

challenging. Previously published STI models have accounted for age mixing in various, suboptimal ways, often due to lacking data. Garnett and Anderson (1993) model age mixing using an assortativity constant, expressing the extent by which people have sexual contacts with partners of their own age only (assortative mixing), or whether people's sexual contacts are distributed across other age groups in proportion to the partner change rates in those age groups (proportionate mixing). Because this approach is simple and flexible it has been used in many STI transmission models, with few attempts to improve this method (Walker et al., 2012). Published models have either used fixed values for the assortativity constant or considered ranges of values (Barnabas et al., 2006; Choi et al., 2010). However, because this constant only summarizes mixing patterns (assuming that the assortativity constant itself does not depend on age), the computed patterns can differ from true mixing patterns. In other models, data about partner ages have been used indirectly (Kretzschmar et al., 1996), or directly but with simplifying assumptions e.g. the same distribution of age differences between sexual partners across all age groups (Low et al., 2007). Turner et al. (2006) used a matrix of probabilities for new partnerships by age combinations, but it is unclear how they obtained these probabilities and balanced the data for men and women. Other models rely on partner age data from one gender only (Regan et al., 2008). Bogaards et al. (2010), in a human papillomavirus transmission model, used data about partner ages from both genders, balancing the discrepancies in data obtained from men and women. This approach is probably the most accurate, but they did not study how combinations of partner ages exactly affect STI transmission. Conclusions about how age mixing impacts STI transmission are further limited by the uncertainty arising from aggregating different data sources on sexual behavior and STI distribution.

The British National Surveys of Sexual Attitudes and Lifestyles (Natsal, [www.natsal.ac.uk](http://www.natsal.ac.uk)) provide a rich source of data about both sexual behavior and chlamydia infection status. The objectives of the present study were to reconstruct detailed age mixing patterns among women and men, to implement these mixing patterns into a mathematical model of *C. trachomatis* transmission, and to identify which age combinations in new partnerships account for most *C. trachomatis* transmission in Britain.

## 2. Methods

### 2.1. Data about age mixing and chlamydia prevalence

The Natsal surveys are decennial national probability sample surveys of the resident British population addressing sexual attitudes, practices, behaviors and lifestyles through interviewer-administered computer-assisted personal interviewing (CAPI) with computer-assisted self-interviewing (CASI) for the more sensitive questions. The second survey (Natsal-2) completed interviews with 12,110 respondents aged 16–44 years in 1999–2001 (response rate 65.4%) and the third (Natsal-3, 2010–2012) included 15,162 respondents aged 16–74 years (response rate 57.7%) (Johnson et al., 2001; Mercer et al., 2013). We used respondent data from Natsal-2 and Natsal-3 on the age at first heterosexual intercourse, number of new heterosexual partners in the year before the interview, and respondent age and partner-age at first sexual intercourse together, with the first, second, and third most recent heterosexual partners. The youngest respondents in the Natsal surveys were 16 years old at the time of the surveys but questions were asked about sexual intercourse from age 13. For the age mixing analyses, we modelled partnerships across a 50 year span, from 13 to 62 years of age. Data from Natsal-2 and Natsal-3 about age mixing were aggregated to increase statistical power. This was warranted because there were no marked differences for the mentioned variables between the surveys for 16–44 year olds, who were represented in both surveys (see Supplementary Information, part A).

We used data about chlamydia infection status to parameterize the

model. Infection status was determined from urine samples for a subsample of 3608 Natsal-2 respondents aged 18–44 years. In the chlamydia transmission model, chlamydia prevalence was assumed to have approached equilibrium. Therefore, we did not use data on chlamydia infection status from Natsal-3: the dynamics of *C. trachomatis* transmission could have been affected by the introduction of a National Chlamydia Screening Programme (NCSP) in the period between the two surveys (PHE, 2016). This program, aimed at people aged 25 years and younger, increased chlamydia testing as it was rolled out across England from 2003 onwards and might have affected prevalence.

### 2.2. Sexual behavior variables contributing to age mixing

We extracted the following variables from Natsal-2 and Natsal-3: the number of new heterosexual partners per year, the proportion of the population that is sexually experienced (ever had sexual intercourse from age 13), and partner ages. We specify the variables for people of sex  $k$  ( $1 = \text{man}/2 = \text{woman}$ ), sexual activity class  $j$  ( $1 = \text{low}/2 = \text{high}$ , defined by number of new sex partners), and age  $a$  (50 one-year age bands from 13 to 62 years), referring to them as being in class  $(k, j, a)$  and use these letters as subscripts in the mathematical notation. A brief description of the fitting and smoothing of variables is given below; for details we refer to the Supplementary Information, part A.

Partner change rates in class  $(k, j, a)$ ,  $c_{k,j,a}$  are defined as the mean number of new heterosexual partners per year for sexually experienced people. We first retrieved the reported number of new partners per year for respondents in class  $(k, a)$  from Natsal data. To account for the variability in the empirical data, we assumed that individuals could belong to one of two sexual activity classes. We estimated  $c_{k,j,a}$  by assuming that the distribution of reported number of new heterosexual partners is the sum of two Poisson distributions with means  $c_{k,j,a}$ , weighted by the proportion of individuals in each sexual activity class,  $f_{k,j}$  (with  $f_{k,2} = 1 - f_{k,1}$ ) (Althaus et al., 2012). Fitting to the data was performed with maximum likelihood estimation methods. Sampling uncertainty across one-year age bands introduces some irregularity in the partner-change rate estimates, so the  $c_{k,j,a}$  values were smoothed across ages  $a$  using a smoothing spline.

To remove sampling irregularity, the heterosexually experienced proportion of the population ( $Q_{k,a}$ ) was described using a Gompertz function, which is here a monotonic increasing function of age. Its parameters were obtained using maximum likelihood estimation with Natsal data on the respondent's age at first heterosexual intercourse.

We constructed  $\rho$  (partner age matrix) as a measure for partner age choice. Matrix entry  $\rho_{k,a,a'}$  is the probability that a sexually active person in class  $(k, a)$  has a new heterosexual partner of age  $a'$ . First, we constructed empirical partner age distributions using data about the respondent's age at first sex with each of their three most recent partners and the ages of those partners. We first fit skew-normal distributions to the bootstrapped empirical distribution of partner ages  $a'$  of each respondent age  $a$  and sex  $k$  by maximum likelihood methods. This resulted in a list of best-fit distribution moments (mean, standard deviation and skewness) for each  $(k, a)$  pair. The best-fit moments showed irregularities across respondent age  $a$  due to sampling uncertainty (Fig. S1). To further smooth the partner age data across respondent ages  $a$ , we performed a linear regression on the means, a second-order polynomial regression on the standard deviations and a logistic regression on the skewness values across respondent ages. The smoothed partner age matrix entry  $\rho_{k,a,a'}$  is given by the density of a skew-normal at age  $a'$  with the corresponding regression values at  $a$  for the three distribution moments.

### 2.3. Describing age mixing patterns derived from sexual behavior variables

We combined the age mixing variables into one composite variable  $T$ , which describes the age mixing patterns in a population while fixing the total supply of and demand for sexual partnerships, using methods

**Table 1**  
Parameters used in the *C. trachomatis* transmission model.

Parameter	Description	Value/Distribution	Source/method
$c_{k,j,a}$	Partner-change per year	Fig. 1A	Natsal-2 and –3
$Q_{k,a}$	Proportion sexually active	Fig. 1B	Natsal-2 and –3
$\rho_{k,a}$	Partner-age distribution	Fig. 1C and D	Natsal-2 and –3
$f_{1,1}$	Proportion men in low activity class	0.94	MLE estimate, see text
$f_{2,1}$	Proportion women in low activity class	0.93	MLE estimate, see text
$q_{k,a}$	Age-dependent probability to become sexually active	$Q_{k,a}/Q_{k,a-1}$	–
$m$	Switching rate sexual activity classes	1	(Fingerhuth et al., 2016)
$\alpha$	Aging rate per year	1	–
$\varepsilon$	Assortativity index between low and high activity class	Prior: U[0,1] Posterior: 0.76 [0.52, 0.99] <sup>a</sup>	MCMC sampling
$\beta$	Per partnership transmission probability	Prior: U[0,1] Posterior: 0.52 [0.47, 0.56] <sup>a</sup>	MCMC sampling
$\gamma$	Infection clearance rate per year	1	(Davies et al., 2014)
$\theta$	Balancing constant	0.5	(Garnett and Anderson, 1994)

<sup>a</sup> Posterior distributions are summarized by posterior means and 95% Bayesian credible intervals.

proposed by (Garnett and Anderson, 1993, 1994).  $T_{k,jj',a,a'}$  quantifies the expected number of new heterosexual partnerships per year between individuals in class  $(k, j, a)$  and class  $(k', j', a')$ :

$$T_{k,jj',a,a'} = N_{k,j,a} \rho_{k,jj',a,a'} c_{k,jj',a,a'} \quad (1)$$

$N_{k,j,a}$  is the expected number of people in class  $(k, j, a)$  who are sexually active. It is calculated by multiplying  $Q_{k,a}$  by  $f_{k,j}$  and by a scaling factor  $M$ , equaling the population size in class  $(k, a)$ . In our transmission model, we track the expected number of heterosexual partnerships in a virtual population of 100,000 that is assumed to be evenly spread across the two sexes and 50 age groups (13–62), i.e.  $M = 1000$  in each group.

$\rho_{k,jj',a,a'}$  is the probability of a sexually active person in class  $(k, j, a)$  having a new sexual partnership with a partner in class  $(k', j', a')$ . It is calculated assuming that the degree of mixing between activity classes is in between assortative (assortativity constant  $\varepsilon = 1$ ) and proportional ( $\varepsilon = 0$ ) (Garnett and Anderson, 1993):

$$\rho_{k,jj',a,a'} = \rho_{k,a,a'} \left( \varepsilon \delta_{jj'} + (1 - \varepsilon) \frac{N_{k',j',a'} c_{k',j',a'}}{\sum_{i=1}^{n_j} N_{k',i,a'} c_{k',i,a'}} \right) \quad (2)$$

Here the Kronecker delta function is defined as  $\delta_{jj'} = 1$  if  $j = j'$  and zero otherwise. We estimated plausible values for  $\varepsilon$  from data using the transmission model (see below).

$c_{k,jj',a,a'}$  represents the adjusted partner change rates, which balance data discrepancies between men and women in the sexually active proportion of the population of a certain age, the partner change rates for that age and the partner ages. To achieve appropriate balancing, we introduce a balancing constant  $B_{jj',a,a'}$ , which is the ratio of new sexual partners per year reported by men (numerator) and women (denominator) for each age and activity class (Garnett and Anderson, 1994):

$$B_{jj',a,a'} = \frac{N_{k,j,a} \rho_{k,jj',a,a'} c_{k,j,a}}{N_{k',j',a'} \rho_{k',jj',a,a'} c_{k',j',a'}} \quad (3)$$

An adjusted set of change rates of sexual partners provides the necessary balancing:

$$c_{k,jj',a,a'} = c_{k,j,a} B_{jj',a,a'}^\theta \quad (4)$$

Here,  $\theta \in [0,1]$  is a constant that weights the data between men and women. Equal weight is given to both sexes when  $\theta = 0.5$  (Garnett and Anderson, 1994). The supply of and demand for sexual contacts for men and women, the numerator and denominator in Eq. (3), respectively, are equal if partner change rates are substituted with the adjusted partner change rates.

#### 2.4. *C. trachomatis* transmission model

We devised a population-based transmission model that integrates the obtained age mixing matrix. We used the model to describe sex- and age-specific chlamydia prevalence. The continuous-time model describes the disease dynamics in a population stratified by sex, age and sexual activity by means of a system of ordinary differential equations:

$$\frac{dU_{k,a}}{dt} = \alpha q_{k,a} U_{k,a-1} - \alpha U_{k,a} \quad (5)$$

$$\begin{aligned} \frac{dS_{k,j,a}}{dt} = & f_{k,j} \alpha (1 - q_{k,a}) U_{k,a-1} + \alpha S_{k,j,a-1} - \alpha S_{k,j,a} - \lambda_{k,j,a} S_{k,j,a} \\ & + \gamma I_{k,j,a} + m f_j \sum_i S_{k,i,a} - m S_{k,j,a} \end{aligned} \quad (6)$$

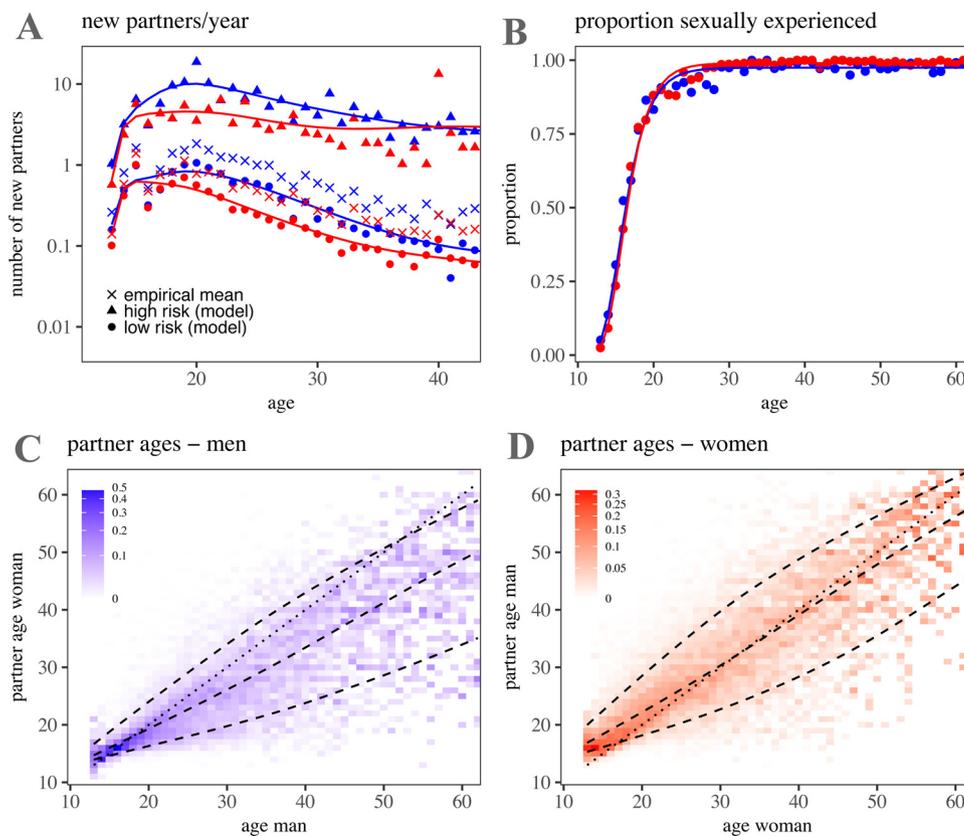
$$\frac{dI_{k,j,a}}{dt} = \alpha I_{k,j,a-1} - \alpha I_{k,j,a} + \lambda_{k,j,a} S_{k,j,a} - \gamma I_{k,j,a} + m f_j \sum_i I_{k,i,a} - m I_{k,j,a} \quad (7)$$

The model assumes two compartments ( $S$ , susceptible; and  $I$ , infectious) for the sexually active population, and one compartment for individuals who have not yet had sexual intercourse ( $U$ ). Aging, at rate  $\alpha$ , moves individuals from one age group to the next. Every year, individuals not having had sex before have a probability  $q_{k,a} = Q_{k,a}/Q_{k,a-1}$  of becoming sexually active depending on their age, and move to either the low or high activity class with probabilities  $f_{k,1}$  and  $f_{k,2}$  respectively. Susceptible individuals can be infected by infectious individuals at rate  $\lambda_{k,j,a}$ , the force of infection:

$$\lambda_{k,j,a} = \beta \sum_{j',a'} c_{k,jj',a,a'} \rho_{k,jj',a,a'} \frac{I_{k',j',a'}(t)}{N_{k',j',a'}(t)} \quad (8)$$

$\beta$  is the per partnership transmission probability. Clearance occurs at rate  $\gamma$ . Individuals leave their sexual activity classes at a rate  $m$  per year, and are redistributed across activity classes with a probability equal to the relative size of the target activity class (Fingerhuth et al., 2016). The total population size remains constant: people leave the oldest age class and are replaced by individuals who enter the youngest age class. Parameter values for the model are shown in Table 1. Data about chlamydia infection status from Natsal-2 were used to estimate  $\beta$  and  $\varepsilon$ , using Markov Chain Monte Carlo (MCMC) sampling (10,000 steps, Supplementary Information, part B) (Chib and Greenberg, 1995).

We computed age- and sex-specific chlamydia prevalences from the model simulations and used them to calculate: i) the expected number of new heterosexual infection-transmitting partnerships per year that involve people of ages  $a$  and  $a'$ , and ii) the probability that a randomly chosen, sexually active person of age  $a$  would be infected with *C. trachomatis* by a new sex-partner of age  $a'$  (Supplementary Information, part C). Using this



**Fig. 1.** Empirical age mixing data, and parametric/smoothed descriptions of these data. **A:** mean number of new sexual partnerships per year (crosses) and partner-change rates ( $c_{k,j,a}$ ) in low (dots) and high (triangles) sexual activity classes. Lines represent smoothed partner-change rates. **B:** proportion of the population that is sexually active per sex and per age ( $Q_{k,a}$ ). Lines represent smoothed proportions. **C, D:** empirical partner-age distributions per respondent age, shown as heat maps, and percentiles (0.10, 0.50 and 0.90) of the fitted and parametrized partner-age matrix ( $\rho_{k,a,a'}$ ), for men (C) and women (D). The values within each partner-age distribution sum to one, to reflect probabilities.

last estimate we computed the odds of being infected with *C. trachomatis* by new partners of a different age, compared with partners of the same age.

### 3. Results

#### 3.1. Age mixing data and age mixing patterns

In the dataset combining Natsal-2 and Natsal-3, sexually experienced men reported more new heterosexual partners per year than sexually experienced women in Britain across all ages (men: mean 0.6, median 0, IQR (0, 1); women: mean 0.3, median 0, IQR (0, 0), Fig. 1A). From the raw data we computed partner change rates. For men, the partner change rates peaked around age 19–20 at 0.8 and 10.1 new sex-partners per year in the low and high activity classes, respectively. For women they peaked around ages 15–19 at 0.6 and 4.6 new sex-partners per year in the low and high activity classes, respectively. In these calculations, 94% and 93% of all men and women, respectively, were attributed to the low sexual activity class and the remainder to the high activity class.

The proportion of sexually experienced people increased sharply after age 13 for both sexes (Fig. 1B). Men and women reported their first experience of heterosexual intercourse at a similar age (men: median 17, IQR (15, 18) years age; women: median 17, IQR (16, 19) years age).

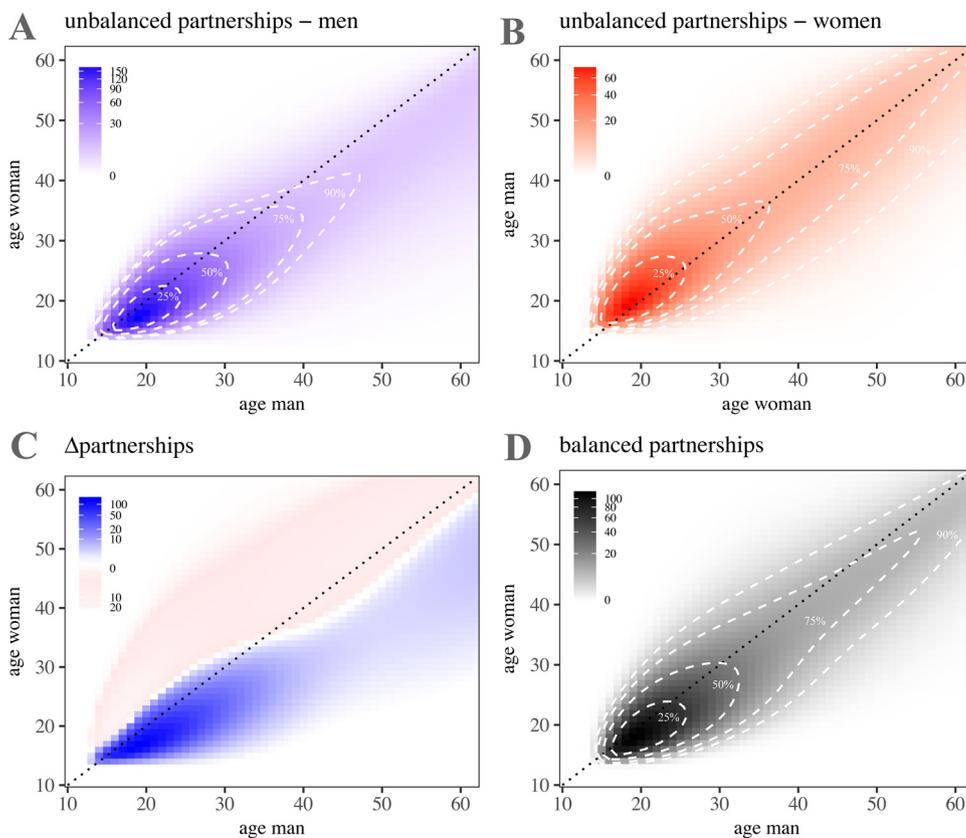
Across all ages considered in this study (13–62 years), men reported younger new female sex partners at the age when they first had sexual intercourse together (median age difference  $\Delta = 4$ , IQR (0, 10) years; age differences from here onwards always defined as age man minus age woman) while women reported that their new sex-partners were of a similar age ( $\Delta = 0$ , IQR (–5, 4) years, Fig. 1C, D). These values show significant variation with respondent age. Men of 16 years and younger reported new sex partners who were older ( $\Delta = -1$ , IQR (–2, 0) years). Men above 16 reported younger new sex partners ( $\Delta = 5$ , IQR (1, 11) years) with the median difference in partner-age increasing with

age. In contrast, women reported new sex-partners that were older from age 13 until age 30 ( $\Delta = 2$ , IQR (0, 5) years), after which their new partners were mostly younger (median  $-2$ , IQR (–7, 3) years). For both sexes, the variance in reported partner ages increases with age. Fitting the partner age data to skew-normal distributions showed a distinct right-skew in these distributions for respondents under 30 (Fig. S1).

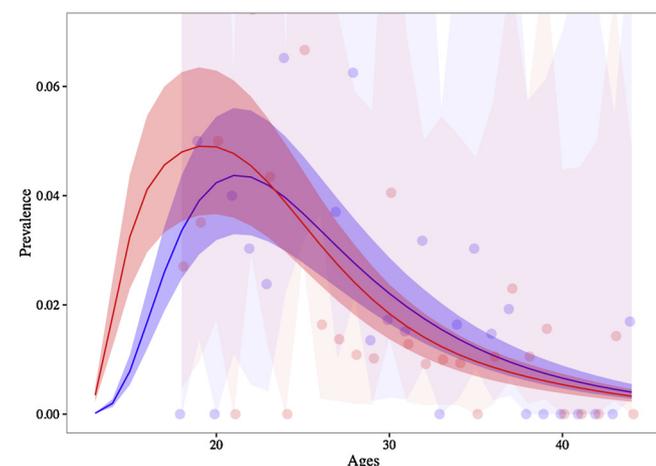
The distributions of reported partner ages of men and women are not consistent with each other. Fig. 2A–C show the degree of mismatch. The balancing procedure, using Eq. (4), generates male and female data that are internally consistent (Fig. 2D). In the balanced data, across all ages, new partnerships tend to involve men who are older than women ( $\Delta = 2$ , IQR (–1, 5) years).

#### 3.2. Impact of age mixing on *C. trachomatis* transmission

The transmission model, incorporating the balanced partnership structure, predicts that chlamydia prevalence rises earlier in young women than in young men, with the peaks at age 19 in women (4.9%, 95% CI 3.6–6.3) and age 21 in men (4.3%, 95% CI 3.2–5.6) (Fig. 3). We used the modelled age-specific chlamydia prevalences to compute the expected number of new partnerships in which *C. trachomatis* is transmitted in a population of 100,000. These numbers are shown as different shades of grey in the heat map in Fig. 4. Each square in the heat map represents one combination of partner ages, and is divided into two triangles: the lower-right triangle shows the number of transmission events from men to women and the upper-left triangle shows the number of transmission events from women to men. By comparing the grey shading between triangles, we can assess the most likely direction of infection transmission for a given partner age combination. For example, *C. trachomatis* transmission is more common from the woman to the man in partnerships in which the woman is 18 and her new male partner is younger than 19 years old. However, if an 18 years old woman gets a new partner of age between 20 and 23, the most likely



**Fig. 2.** The total number of new heterosexual partnerships per year for each combination of ages in a population of 100,000, shown as heat maps. **A:** Unbalanced partnerships for men. **B:** Unbalanced partnerships for women. **C:** Differences in unbalanced partnerships between men and women. **D:** Balanced partnerships (matrix T, with the activity classes aggregated within the same sex). The circular contour lines in A, B and D (from smaller to larger) represent 25%, 50%, 75% and 90% of values inside the contours.



**Fig. 3.** Estimated age-specific prevalence of chlamydia, for men (blue) and women (red) and 95% CI, as predicted by the model. Dots show observed chlamydia prevalence in one-year age bands from Natsal-2, including contours of 95% binomial confidence intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

direction of infection transmission is from the man to the woman; if this new partner is 24 or older, the most likely direction of infection transmission is again from the woman to the man.

Incident *C. trachomatis* infections are more likely to be generated in new sexual partnerships of younger, rather than older people. Roughly fifty percent of incident infections are generated in new partnerships when men are between 16 and 28 and women are between 16 and 26 years old, with an age difference not exceeding four years. However, in > 50% of infection transmitting partnerships, at least one partner was 25 years or older.

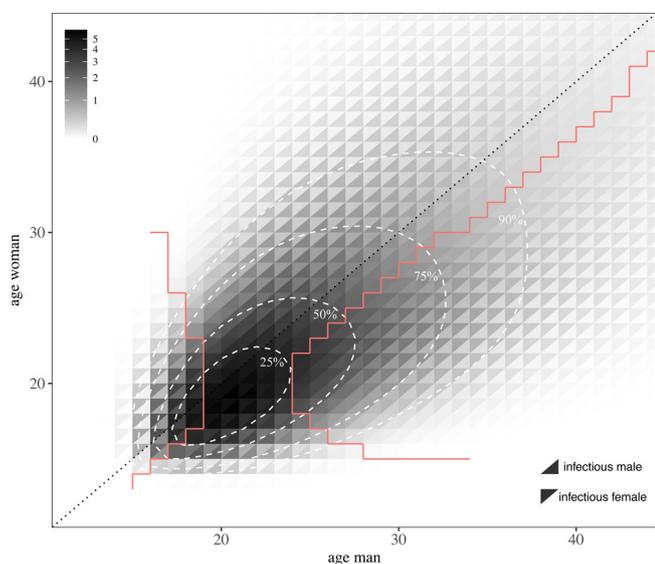
Only under the age of 20 do men and women have higher odds of

becoming infected with *C. trachomatis* in a new sexual partnership with an older partner compared with a partner of the same age; above that age the odds are reversed (Fig. 5). In contrast, people older than 22 years have higher odds of becoming infected by a new heterosexual partner more than two years younger compared with a partner of same age. There are no major differences between men and women in these age-dependent odds.

#### 4. Discussion

Our analysis of age mixing in new heterosexual partnerships in Britain reveals pronounced discrepancies in reporting of partner age differences and in the number of new heterosexual partners per year reported by men and women (Fig. 1). In a transmission dynamic model with a balanced partnership structure, we showed at which age combinations new *C. trachomatis* infections are generated and in which direction transmission occurs (Fig. 4). > 50% of incident chlamydia infections occur in age groups older than 25 years.

Our approach to reconciling age mixing data from men and women is useful in the majority of cross-sectional surveys that do not collect data from both members of a sexual partnership. Reconstructing sexual network properties from egocentric data is difficult, particularly in the presence of bias (Britton and Trapman, 2015). The sex-specific differences in reports of age mixing need to be balanced before they can be incorporated into an internally consistent model of STI transmission. We integrated several independent aspects of age mixing into one consolidated structure and obtained a complete representation of the mixing patterns, combining and balancing the male and female age mixing behavior. The use of a balancing constant to combine the age mixing variables is a standard procedure in mathematical modeling studies (Garnett and Anderson, 1994), but the choice of  $\theta$  is subjective. We used a value of  $\theta = 0.5$ , which puts the balanced age mixing matrix in between its equivalent unbalanced matrices while giving equal weight to data reported by women and men. Some modeling studies



**Fig. 4.** Expected number of yearly heterosexual partnerships, in which *C. trachomatis* is transmitted between sexual partners for different age combinations in a population of 100,000 and transmission direction. Every combination of ages is represented as a square in the heat map, and comprises two triangles. Lower-right triangle (▲): transmission events from men to women. Upper-left triangle (▼): transmission events from women to men. The shading of each triangle represents the expected numbers of new infections per year in a population of 100,000 individuals. A darker-shaded upper-left triangle indicates more women to men transmissions, while the opposite is true for a darker-shaded lower-right triangle. The sum of the densities of two triangles determines the expected total number of *C. trachomatis* infections for one combination of ages. The circular contour lines represent percentage of values inside the contours (summed number of transmission events for men and women). The red lines demarcate three regions of partner age combinations in which a particular direction of infection transmission is likely. In the regions to the left and right, infections are most likely to be transmitted from women to men. In the region in the middle of the plot, infections are most likely to be transmitted from men to women. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

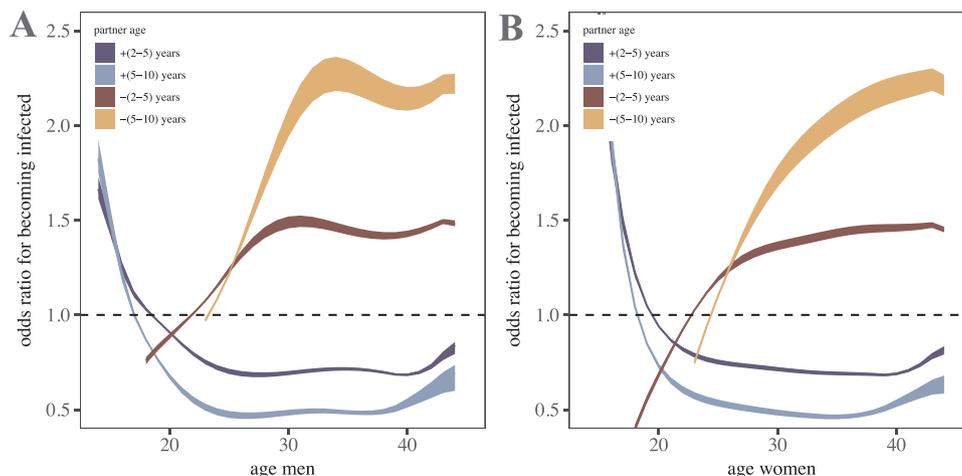
have used data from one sex only (Regan et al., 2008) which comes with substantial challenges when interpreting the results. Relying on data from men only ( $\theta = 0$ ) would increase the estimated partner change rates as well as the proportion of age disparate relationships (Fig. 2A). Therefore, data from men only will increase the estimated *C. trachomatis* prevalences when leaving transmission parameters

unchanged; relying on data from women only ( $\theta = 1$ ) does the opposite. If asymmetric data on sexual behavior of men and women stem from social desirability bias (Fenton et al., 2001), then further research on the magnitude of the bias may lead to improved balancing values.

The probability of acquiring a *C. trachomatis* infection in age disparate partnerships depends on the age difference as well as the age of each partner. This is evidenced by the direction of infection in the model, but not always considered when studying the effects of age disparate partnerships on STI prevalence. For example, the Malawian study used aggregated data from men and women between 18 and 49 years (Beauclair et al., 2016). Their results suggested that women with either a 5–10 year older or a 5–10 year younger sexual partner tended to have a higher risk of HIV infection. Our study can also help explain why Stein et al. (2008) failed to find higher odds of *C. trachomatis* prevalence in women between 18 and 26 years who had sex with male partners 2–5 years older compared with those who had partners of a similar age. As shown in Fig. 5, this is the likely outcome of aggregating data from women both below and above 20, where the odds ratios undergo the most pronounced changes, and thus dampening epidemiological signal.

The Natsal datasets have several advantages over other data sources for the investigation of patterns of sexual mixing by age. To our knowledge, few cross-sectional probability surveys are large and detailed enough to provide precise data about sexual behavior for men and women across a wide age range (Wellings et al., 2006). Natsal also collects urine specimens so information about *C. trachomatis* infection status came from the same population as the behavioral data, although the number of specimens was smaller so the estimates of chlamydia prevalence in single age bands was rather imprecise. These data allowed us to reconstruct the age mixing structure of new heterosexual partnerships in a geographically defined region with unprecedented resolution, whilst using established mathematical methods (Garnett and Anderson, 1993, 1994) and to investigate how age mixing affects chlamydia prevalence and incidence.

As with all models, our approach has some limitations. Our aim was to eliminate model complexity to such a degree that the effects of age mixing could be studied in isolation while still capturing essential features of *C. trachomatis* behavior. All parameters not specifying age mixing ( $\beta$ ,  $\gamma$ ,  $m$ ) were chosen to be identical among sexes and across ages. Sex- and age-specific differences, for example, in factors that affect transmissibility and acquired immunity could also influence our results (Geisler et al., 2013). Also, we assumed no chlamydia testing in the transmission model. We calibrated the model to chlamydia prevalence data collected in 1999–2000. At that time, most testing took place in specialized STI clinics. Chlamydia testing rates did not increase



**Fig. 5.** Odds ratios for men (A) and women (B) of different ages for becoming infected by new heterosexual partners of a discordant age compared to new heterosexual partners of the same age ( $\pm 1$  year), for different discordant age ranges (positive if the partner is older; negative otherwise).

markedly until the NCSP in England promoted widespread testing of sexually active women and men between 15 and 24 years and began its roll-out in 2003 (Chandra et al., 2017).

The methodological approaches proposed here can be used for more detailed study of the effects of age mixing on the transmission dynamics of *C. trachomatis* and other STIs. For example, our model predicts that over 50% of new chlamydia infections are generated outside of the target age group of chlamydia testing interventions in most countries, including the NCSP in England. Additional data about chlamydia testing in different age groups in women and men over time would help to determine the relative effectiveness of programs targeted at different age groups. The duration of HIV infection is much longer than for chlamydia so and modeling could show whether the age combinations at which the greatest number of infection transmitting partnerships occur differs from those for *C. trachomatis* (CDC, 2015). In summary, we conducted a detailed analysis of patterns of age mixing obtained from population survey data, balancing inconsistencies of sex-specific reports. In a mathematical model, we found that the direction of transmission of *C. trachomatis* differs according to the age combinations of sexual partners. Our method of incorporating age mixing data into dynamic STI transmission models can be used to study the age-dependent transmission of any STI and to understand the effects of partner-age differences on STI incidence.

### Funding statement

JHS and VG received support from the Swiss National Science Foundation (grant number 160320). CLA received support from the LUSTRUM Program of Research, funded by the National Institute for Health Research (NIHR) under its Program Grants for Applied Research Program (Reference Number RP-PG-0614-20009). Natsal-3 was supported by grants from the UK Medical Research Council ([www.mrc.ac.uk](http://www.mrc.ac.uk); grant number G0701757; funding source for Natsal-3) and the Wellcome Trust ([www.wellcome.ac.uk](http://www.wellcome.ac.uk); 084840; funding source for Natsal-3), with contributions from the Economic and Social Research Council and Department of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The views expressed are those of the author(s) and not necessarily those of the funders.

### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.epidem.2018.03.004>.

### References

- 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, e51438.
- Anderson, R.M., May, R.M., Ng, T.W., Rowley, J.T., 1992. Age-dependent choice of sexual partners and the transmission dynamics of HIV in Sub-Saharan Africa. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 336, 135–155.
- Bajos, N., Wadsworth, J., Ducot, B., Johnson, A.M., Le Pont, F., Wellings, K., Spira, A., Field, J., 1995. Sexual behaviour and HIV epidemiology: comparative analysis in France and Britain. *ACSF Group. AIDS* 9, 735–743.
- Barnabas, R.V., Laukkanen, P., Koskela, P., Kontula, O., Lehtinen, M., Garnett, G.P., 2006. Epidemiology of HPV 16 and cervical cancer in Finland and the potential impact of vaccination: mathematical modelling analyses. *PLoS Med.* 3, e138.
- Beauchair, R., HELLERINGER, S., HENS, N., DELVA, W., 2016. Age differences between sexual partners, behavioural and demographic correlates, and HIV infection on Likoma Island, Malawi. *Sci. Rep.* 6, 36121.
- Bogaards, J.A., Xiridou, M., Coupe, V.M., Meijer, C.J., Wallinga, J., Berkhof, J., 2010. Model-based estimation of viral transmissibility and infection-induced resistance from the age-dependent prevalence of infection for 14 high-risk types of human papillomavirus. *Am. J. Epidemiol.* 171, 817–825.
- Britton, T., Trapman, P., 2015. Inferring global network properties from egocentric data with applications to epidemics. *Math. Med. Biol.* 32, 99–112.
- CDC, 2015. HIV Surveillance Report. Centers for Disease Control and Prevention (CDC).
- Chandra, N.L., Soldan, K., Dangerfield, C., Sile, B., Duffell, S., Talebi, A., Choi, Y.H., Hughes, G., Woodhall, S.C., 2017. Filling in the gaps: estimating numbers of chlamydia tests and diagnoses by age group and sex before and during the implementation of the English National Screening Programme, 2000–2012. *Euro Surveill.* 22.
- Chib, S., Greenberg, E., 1995. Understanding the metropolis-hastings algorithm. *Am. Stat.* 49, 327–335.
- Choi, Y.H., Jit, M., Gay, N., Cox, A., Garnett, G.P., Edmunds, W.J., 2010. Transmission dynamic modelling of the impact of human papillomavirus vaccination in the United Kingdom. *Vaccine* 28, 4091–4102.
- Davies, B., Anderson, S.J., Turner, K.M., 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.
- Fenton, K.A., Johnson, A.M., McManus, S., Erens, B., 2001. Measuring sexual behaviour: methodological challenges in survey research. *Sex. Transm. Infect.* 77, 84–92.
- Fingerhuth, S.M., Bonhoeffer, S., Low, N., Althaus, C.L., 2016. Antibiotic-Resistant neisseria gonorrhoeae spread faster with more treatment, not more sexual partners. *PLoS Pathog.* 12, e1005611.
- Ford, K., Sohn, W., Lepkowski, J., 2001. Characteristics of adolescents' sexual partners and their association with use of condoms and other contraceptive methods. *Fam. Plann. Perspect.* 33, 100–105 (132).
- 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, 137–159.
- Garnett, G.P., Anderson, R.M., 1994. Balancing sexual partnerships in an age and activity stratified model of HIV transmission in heterosexual populations. *IMA J. Math. Appl. Med. Biol.* 11, 161–192.
- Garnett, G.P., Anderson, R.M., 1996. Sexually transmitted diseases and sexual behavior: insights from mathematical models. *J. Infect. Dis.* 174 (Suppl. 2), S150–S161.
- Geisler, W.M., Lensing, S.Y., Press, C.G., Hook 3rd, E.W., 2013. Spontaneous resolution of genital Chlamydia trachomatis infection in women and protection from reinfection. *J. Infect. Dis.* 207, 1850–1856.
- Gregson, S., Nyamukapa, C.A., Garnett, G.P., Mason, P.R., Zhuwau, T., Carael, M., Chandiwana, S.K., Anderson, R.M., 2002. Sexual mixing patterns and sex-differentials in teenage exposure to HIV infection in rural Zimbabwe. *Lancet* 359, 1896–1903.
- Helleringer, S., Kohler, H.P., Mkandawire, J., 2011. Women underestimate the age of their partners during survey interviews: implications for HIV risk associated with age mixing in northern Malawi. *Sex. Transm. Dis.* 38, 1030–1035.
- Johnson, A.M., Mercer, C.H., Erens, B., Copas, A.J., McManus, S., Wellings, K., Fenton, K.A., Korovessis, C., Macdowall, W., Nanchahal, K., Purdon, S., Field, J., 2001. Sexual behaviour in Britain: partnerships, practices, and HIV risk behaviours. *Lancet* 358, 1835–1842.
- Kraut-Becher, J.R., Aral, S.O., 2006. Patterns of age mixing and sexually transmitted infections. *Int. J. STD AIDS* 17, 378–383.
- Kretzschmar, M., van Duynhoven, Y.T., Severijnen, A.J., 1996. Modeling prevention strategies for gonorrhoea and Chlamydia using stochastic network simulations. *Am. J. Epidemiol.* 144, 306–317.
- Leclerc-Madlala, S., 2008. Age-disparate and intergenerational sex in southern Africa: the dynamics of hypervulnerability. *AIDS* 22 (Suppl. 4), S17–25.
- Low, N., McCarthy, A., Macleod, J., Salisbury, C., Campbell, R., Roberts, T.E., Horner, P., Skidmore, S., Sterne, J.A., Sanford, E., Ibrahim, F., Holloway, A., Patel, R., Barton, P.M., Robinson, S.M., Mills, N., Graham, A., Herring, A., Caul, E.O., Davey Smith, G., Hobbs, F.D., Ross, J.D., Egger, M., Chlamydia Screening Studies Project G, 2007. Epidemiological, social, diagnostic and economic evaluation of population screening for genital chlamydial infection. *Health Technol. Assess.* 11 (iii-iv ix-xii), 1–165.
- Mercer, C.H., Copas, A.J., Sonnenberg, P., Johnson, A.M., McManus, S., Erens, B., Cassell, J.A., 2009. Who has sex with whom? Characteristics of heterosexual partnerships reported in a national probability survey and implications for STI risk. *Int. J. Epidemiol.* 38, 206–214.
- Mercer, C.H., Tanton, C., Prah, P., Erens, B., Sonnenberg, P., Clifton, S., Macdowall, W., Lewis, R., Field, N., Datta, J., Copas, A.J., Phelps, A., Wellings, K., Johnson, A.M., 2013. Changes in sexual attitudes and lifestyles in Britain through the life course and over time: findings from the National Surveys of Sexual Attitudes and Lifestyles (Natsal). *Lancet* 382, 1781–1794.
- Ott, M.Q., Barnighausen, T., Tanser, F., Lurie, M.N., Newell, M.L., 2011. Age-gaps in sexual partnerships: seeing beyond 'sugar daddies'. *AIDS* 25, 861–863.
- PHE, 2016. National Chlamydia Screening Programme (NCSP): Data Tables.
- 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.
- Stein, C.R., Kaufman, J.S., Ford, C.A., Feldblum, P.J., Leone, P.A., Miller, W.C., 2008. Partner age difference and prevalence of chlamydial infection among young adult women. *Sex. Transm. Dis.* 35, 447–452.
- Swartzendruber, A., Zenilman, J.M., Nicolai, L.M., Kershaw, T.S., Brown, J.L., Diclemente, R.J., Sales, J.M., 2013. It takes 2: partner attributes associated with sexually transmitted infections among adolescents. *Sex. Transm. Dis.* 40, 372–378.
- Turner, K.M., Adams, E.J., Gay, N., Ghani, A.C., Mercer, C., Edmunds, W.J., 2006. Developing a realistic sexual network model of chlamydia transmission in Britain. *Theor. Biol. Med. Model.* 3, 3.
- UNAIDS, 2016. AIDS by the Numbers.
- Walker, R., Nickson, C., Lew, J.B., Smith, M., Canfell, K., 2012. A revision of sexual mixing matrices in models of sexually transmitted infection. *Stat. Med.* 31, 3419–3432.
- Wellings, K., Collumbien, M., Slaymaker, E., Singh, S., Hodges, Z., Patel, D., Bajos, N., 2006. Sexual behaviour in context: a global perspective. *Lancet* 368, 1706–1728.