A metapopulation modelling framework for gonorrhoea and other sexually transmitted infections in heterosexual populations

Mark I. Chen[1,2,*], **Azra C. Ghani**[3] **and W. John Edmunds**[1,2]

[1]*Health Protection Agency, Centre for Infections, London NW9 5EQ, UK*
[2]*London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK*
[3]*Imperial College London, London W2 1PG, UK*

Gonorrhoea continues to be a public health problem in the UK, and is the second most common bacterial sexually transmitted infection (STI) after chlamydia. In the UK, gonorrhoea is disproportionately concentrated in epidemiologically distinct subpopulations, with much higher incidence rates in young people, some ethnic minorities and inner city subpopulations. The original model of STI transmission proposed by Hethcote and Yorke explained some of these features through the concept of the 'core group'. Since then, several authors have modified the original model approach to include multiple sexual activity classes, but found this modelling approach to be inadequate when applied to low-prevalence settings such as the UK. We present a metapopulation framework for modelling gonorrhoea and other STIs. The model proposes that the epidemiology of gonorrhoea is largely driven by subpopulations with higher than average concentrations of individuals with high sexual risk activity. We show how this conceptualization of gonococcal epidemiology overcomes key limitations associated with some of the prior efforts to model gonorrhoea. We also use the model to explain several epidemiological features of gonorrhoea, such as its asymmetric distribution across subpopulations, and the contextual risk experienced by members of at-risk subpopulations. Finally, we extend the model to explain the distribution of other STIs, using chlamydia as an example of a more ubiquitous bacterial STI.

Keywords: mathematical model; epidemiology; gonorrhoea; *Chlamydia trachomatis*; sexual behaviour; sexually transmitted diseases

1. INTRODUCTION

Gonorrhoea is often disproportionately concentrated in epidemiologically distinct subpopulations. For instance, in the UK, incidence rates of gonorrhoea are much higher in young people and some ethnic minorities (Lacey *et al.* 1997; Low *et al.* 1997; Hickman *et al.* 1999), and the disease is characterized by intense spatial clustering (Shahmanesh *et al.* 2000; Monteiro *et al.* 2005; Risley *et al.* 2007), suggesting that some demographically and geographically distinct subpopulations are at much greater risk of infection. Similar observations have been made in the USA, where gonorrhoea and other sexually transmitted infections (STIs) have been shown in multiple studies to be concentrated in certain geographical areas (Rothenberg 1983; Potterat *et al.* 1985; Bernstein *et al.* 2004; Kerani *et al.* 2005). Moreover, black ethnicity and area of residence are often overlapping risk factors both in the USA (Rothenberg & Potterat

1988; Becker *et al.* 1998; Jennings *et al.* 2005) and the UK (Monteiro *et al.* 2005; Risley Ward *et al.* 2007), although one of the USA studies from Baltimore City, Maryland, suggests that the effect of geographical clustering persists after adjusting for the effect of ethnicity (Jennings *et al.* 2005). That study also goes on to suggest that the results 'support the conclusion that risks for gonorrhoea are associated with definable sociogeographic spaces, which may be fundamental ecologic units of STI transmission', and that public health resources in the form of enhanced STI testing and treatment services could be targeted at some of these at-risk areas to interrupt transmission. A better appreciation of the potential impact of such interventions would in turn be facilitated by an improved understanding of the underlying disease transmission dynamics giving rise to the observed sociogeographic epidemiology of gonorrhoea and other STIs.

Models of STI transmission can help us to develop our understanding of a system, and to evaluate the potential impact of proposed interventions (Garnett 2002). The original model of STI transmission proposed by Hethcote & Yorke (1984) suggested that the critical element for the persistence of gonorrhoea is the

*Author and address for correspondence: Modelling and Economics Unit, Centre for Infections, Health Protection Agency, 61 Colindale Avenue, London NW9 5EQ, UK (mark_chen@pacific.net.sg).

existence of a 'core group' with a high rate of change of sexual partners. 'Sexual mixing' between members of the core group then maintains a subpopulation of individuals with a high STI prevalence within a larger community with much lower partner change rates and disease prevalence.

Epidemiological studies subsequently lent support to the core group hypothesis. Geographical locales where gonorrhoea incidence was much higher than surrounding areas have been identified as possible 'core' populations (Rothenberg 1983; Bernstein *et al.* 2004). Moreover, individuals in such locales formed sexual partnerships with each other more frequently than would be expected by chance (Potterat *et al.* 1985; Zenilman *et al.* 1999), suggesting a pattern of 'assortative mixing' between members of core populations. However, when Garnett and colleagues applied a modified version of the model by Hethcote and Yorke to empirical data (Garnett *et al.* 1999), the model suggested that, contrary to the assortative mixing observed in empirical studies, 'near-random mixing' was required to reproduce the observed distribution of gonorrhoea. In addition, Garnett's work suggested that their modelling approach had other limitations in that it was highly sensitive to assumptions on the parameters that describe gonococcal transmission, and as such was inadequate for testing interventions aimed at reducing gonococcal transmission.

There is an ambiguity as to what constitutes the core population for STI transmission. While the original model by Hethcote and Yorke stipulated only two groups with different sexual partner change rates, several later models introduced multiple sexual activity classes (ACs) defined by their average partner change rates (e.g. Garnett *et al.* 1999). These models often formulate their ACs by grouping individuals from population-based surveys of sexual behaviour on the basis of the number of partners reported over an appropriate time period (e.g. Turner *et al.* 2004; White *et al.* 2005). In such models, which we shall henceforth refer to as 'AC models', the dynamics of infection is largely driven by the most sexually active classes, potentially leading to the interpretation that the core group for STI transmission is of individuals with high partner change rates. However, a review by Thomas & Tucker (1996) points out that the term core group has also been used to refer to several other overlapping concepts in the STI literature. Outside of mathematical modelling, the core group has been used to refer to individuals at higher risk of infection such as those with multiple episodes of STIs, or a geographical cluster of cases, which Thomas and Tucker termed the clinical–epidemiologic perspective. In addition, the core group has been used to describe groups of persons, including uninfected individuals, who could be defined socioculturally and might be at higher risk from STIs, such as sex workers and gang members (the sociocultural perspective).

We suggest that what had been empirically demonstrated in most studies was actually the existence of what we would call 'core subpopulations', comprising individuals who were members of some form of sexual network (Wasserheit & Aral 1996). Membership of these subpopulations was not dependent on the intensity of individual risk behaviour but on identifiable geographical and sociodemographic factors. Further evidence for the existence of such sexual networks comes from recent molecular epidemiologic studies in London and Amsterdam, which showed that distinct strains of gonorrhoea caused geographically clustered infections in individuals with significantly different sociodemographic characteristics (Choudhury *et al.* 2006; Kolader *et al.* 2006; Risley *et al.* 2007). We interpret this as further support for the concept that the transmission of gonorrhoea and possibly other STIs is in essence being maintained by geographically and demographically distinct core subpopulations, with such subpopulations comprising not only the cases observed in the studies mentioned (the clinical–epidemiologic perspective), but also the uninfected individuals belonging to the same sexual network, which could be defined by a combination of sociocultural and geographical characteristics (i.e. the sociocultural perspective).

In this work, we propose a metapopulation (MP) modelling framework which we believe more accurately depicts epidemiological observations on gonorrhoea, and STIs in general, than the AC models of STI transmission. Instead of defining a core group based on individual risk, the population is divided into a number of subpopulations of finite size representing the possible sexual networks that may exist within a given community. The sections that follow will

— describe the conceptual and mathematical framework behind the MP model,
— use the example of gonorrhoea, modelled as an susceptible–infectious–susceptible (SIS)-type infection to illustrate the model's key features,
— explore the influence of key parameters for the MP model,
— compare the MP model to the AC model of STI transmission, and
— illustrate how the model can explain epidemiological characteristics of gonorrhoea as well as be generalized to other STIs such as chlamydia, which, as gonorrhoea, is an SIS type bacterial STI, but which is more ubiquitous in its distribution (Shahmanesh *et al.* 2000; Monteiro *et al.* 2005).

Finally, we will discuss the implications of our MP model for understanding the epidemiology of STIs, as well as some agenda for future research.

2. METHODS

2.1. MP framework

The term MP comes from the field of population ecology, and implies a 'population of local populations' (Hanski 1998), i.e. a population composed of several subpopulations that are linked in some way to each other. Here, we use a MP to depict a collection of individuals at risk of infection with a given STI. In its application, the MP would represent all individuals from sexually active age groups resident in some reporting unit commonly used in measuring and aggregating data on sexual behaviour and disease
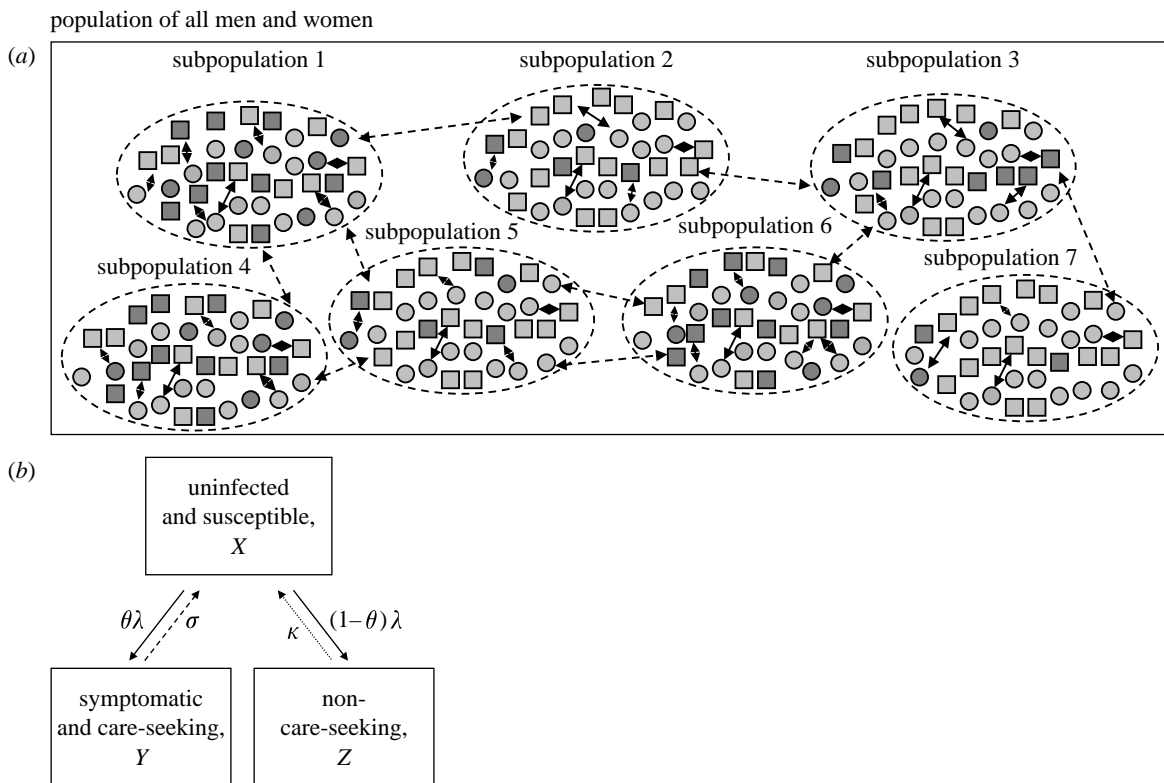
population of all men and women

Figure 1. (*a*) Representation of a MP model for sexually active heterosexual men and women. The heterosexual population comprises subpopulations with an equal number of men and women. The proportion of high-activity individuals is allowed to vary between subpopulations (light grey squares, low-activity men; dark grey squares, high-activity men; light grey circles, low-activity women; dark grey circles, high-activity women; solid two-head arrows, partnerships within subpopulations; dashed two-head arrows, partnerships between subpopulations). (*b*) Schematic of infection states for gonorrhoea. Individuals have three infection states: uninfected and susceptible, symptomatic and care-seeking, and non-care-seeking. Arrows represent transitions between infection states (solid arrows, λ: rate of infection from contact with infectious members of opposite gender; dashed arrows, σ: recovery rate in symptomatic care-seeking patients; dotted arrows, κ: recovery rate in non-care-seeking patients).

incidence, e.g. the Greater London population within the UK, which then could be seen as composed of smaller subpopulations defined by geographical and sociodemographic features.

In our model, the MP is organized into n number of subpopulations of finite size S, each with an equal number of men and women (figure 1*a*). Sexual partnerships are formed more frequently between members of the same subpopulation than between members of different subpopulations. These subpopulations can be thought of as representing the demographic, social, cultural and geographical boundaries, which may restrict the formation of sexual partnerships. The extent to which sexual partnerships are formed with others within the same subpopulation can be defined by letting a proportion of partnerships, π, occur exclusively with members of the opposite gender in the same subpopulation, with the remainder $(1-\pi)$ occurring with available opposite gender partnerships drawn randomly from all subpopulations. We shall refer to the former type of sexual partnerships as 'local sexual partnerships' and the latter as 'global sexual partnerships'. Although it has been postulated that subpopulations may differ in other aspects such as access to healthcare (Wasserheit & Aral 1996), we assume for simplicity that the main difference between subpopulations is in the proportion of individuals who engage in a high level of unprotected sexual risk

activity. To depict this in the most parsimonious fashion, we model just two sexual ACs within each subpopulation—low sexual activity individuals and high sexual activity individuals. This structure creates the possibility that individuals can have two levels of risk, one related to their personal sexual activity and the other to the sexual activity of the subpopulation to which they belong. However, before exploring this, we first describe the STI transmission model.

2.2. AC and MP models of STI transmission

We based our STI transmission model largely on Garnett's modification of the original model by Hethcote & Yorke (1984; Garnett *et al.* 1999). We use the subscripts i for subpopulation membership, j for gender (1 for men and 2 for women) and k for sexual AC (1 for low, and 2 for high sexual activity). Three infection states are modelled (figure 1*b*). Individuals in the susceptible and uninfected state (X) could, upon infection, be symptomatic and seek care (symptomatic care-seeking infections, Y); they may also be asymptomatic or fail to seek care from their symptoms (non-care-seeking infections, Z). Recovery from either of these infectious states returns individuals to the susceptible state at a faster rate (σ_j) for symptomatic care-seeking individuals and a slower rate (κ_j) for non-care-seeking individuals. Partnerships are modelled as

instantaneous events, occurring at rates $c_{j,k}$ for gender j and sexual AC k, with transmission occurring in partnerships between susceptible and infectious members of opposing genders at a fixed per partnership transmission probability (β_j), which may differ by gender (male-to-female versus female-to-male).

The AC models of STI transmission have been described in detail by others (Garnett *et al.* 1999), and we reproduced a simple version of this model for comparison with our MP model. The model had two sexual ACs and a parameter ε, which governs the degree of assortative sexual mixing of the two ACs, where $\varepsilon=1$ denotes fully assortative mixing (i.e. high-activity individuals form partnerships exclusively with other high-activity individuals, and likewise for low-activity individuals) and $\varepsilon=0$ denotes fully random mixing (i.e. sexual mixing is proportionate to the number of available partnerships). The notation used is similar to that which will be used for the MP model, except that the subscript i is not used, since the simplest form of the AC model is only divided into ACs and not into subpopulations. The model can be described by the following equations:

$$\frac{\mathrm{d}X_{j,k}}{\mathrm{d}t} = \sigma_j Y_{j,k} + \kappa_j Z_{j,k} - \lambda_{j,k} X_{j,k}, \qquad (2.1)$$

$$\frac{\mathrm{d}Y_{j,k}}{\mathrm{d}t} = \theta_j \lambda_{j,k} X_{j,k} - \sigma_j Y_{j,k}, \qquad (2.2)$$

$$\frac{\mathrm{d}Z_{j,k}}{\mathrm{d}t} = (1-\theta_j)\lambda_{j,k} X_{j,k} - \kappa_j Z_{j,k}, \qquad (2.3)$$

where θ_j is the proportion of new infections that seek care from symptoms, and $\lambda_{j,k}$ is the force of infection experienced by members of gender j and sexual AC k. The force of infection is determined by the prevalence of infection in opposite gender individuals weighted for the infectious partnerships formed by those of different sexual activity states. In the equations that follow, we use the notation $3-j$ to reference opposite gender individuals (since when $j=1$, $3-j=2$, and when $j=2$, $3-j=1$). The force of infection term has to account for the degree of assortative sexual mixing between ACs, and hence a term for sexual mixing as follows:

$$\lambda_{j,k}=\beta_j c_{j,k}\left(\rho_{j,k,1}\frac{Y_{3-j,1}+Z_{3-j,1}}{P_{3-j,1}Q_{3-j}}+\rho_{j,k,2}\frac{Y_{3-j,2}+Z_{3-j,2}}{P_{3-j,2}Q_{3-j}}\right), \qquad (2.4)$$

where $P_{3-j,1}$ is the proportion in the AC of the opposing gender, and $\rho_{j,k,1}$ and $\rho_{j,k,2}$ are the terms for sexual mixing as a function of ε, and are given as

$$\rho_{j,1,1}=\varepsilon+(1-\varepsilon)c_{3-j,1}P_{3-j,1}Q_{3-j}/N, \qquad (2.5a)$$

$$\rho_{j,1,2}=(1-\varepsilon)c_{3-j,2}P_{3-j,2}Q_{3-j}/N, \qquad (2.5b)$$

$$\rho_{j,2,1}=(1-\varepsilon)c_{3-j,1}P_{3-j,1}Q_{3-j}/N, \qquad (2.5c)$$

$$\rho_{j,2,2}=\varepsilon+(1-\varepsilon)c_{3-j,2}P_{3-j,2}Q_{3-j}/N. \qquad (2.5d)$$

Here, N is the total number of partnerships formed per unit time, and is defined as

$$N=Q_1(c_{1,1}P_{1,1}+c_{1,2}P_{1,2})=Q_2(c_{2,1}P_{2,1}+c_{2,2}P_{2,2}), \qquad (2.6)$$

where Q_1 is the total number of men and Q_2 the total number of women in the population.

The expression for R_0 in the AC model can be subdivided into the terms describing the average number of female infections produced by the 'typical' infectious male, R^{M}, and the average number of male infections produced by the typical infectious female, R^{F}, where

$$R^{\mathrm{M}} = \beta_2 d_1 \begin{pmatrix} c_{1,1}\rho_{1,1,1} & c_{1,1}\rho_{1,1,2} \\ c_{1,2}\rho_{1,2,1} & c_{1,2}\rho_{1,2,2} \end{pmatrix}, \qquad (2.7a)$$

$$R^{\mathrm{F}} = \beta_1 d_2 \begin{pmatrix} c_{2,1}\rho_{2,1,1} & c_{2,1}\rho_{2,1,2} \\ c_{2,2}\rho_{2,2,1} & c_{2,2}\rho_{2,2,2} \end{pmatrix}, \qquad (2.7b)$$

where $\rho_{j,k,l}$ are the terms for sexual mixing previously defined in equation $(2.5a)$–$(2.5d)$, and d_j is the average duration of infectiousness weighted for the proportion of care-seeking and non-care-seeking infections

$$d_j = \theta_j/\sigma_j + (1-\theta_j)/\kappa_j, \qquad (2.8)$$

with β_j, $c_{j,k}$, θ_j, σ_j and κ_j being the respective parameters previously described.

R_0 is the geometric mean of R^{M} and R^{F},

$$R_0 = \sqrt{R^{\mathrm{M}} R^{\mathrm{F}}}. \qquad (2.9)$$

R_0 can thus be computed by first defining, and then finding the dominant eigenvalue for a 'next-generation operator' based on equations $(2.7a)$ and $(2.7b)$ (Heesterbeek 2002).

The MP model can likewise be expressed as a set of differential equations very similar to those used for the AC model, with one key difference being the use of the subscript i to denote subpopulation membership,

$$\frac{\mathrm{d}X_{i,j,k}}{\mathrm{d}t} = \sigma_j Y_{i,j,k} + \kappa_j Z_{i,j,k} - \lambda_{i,j,k} X_{i,j,k}, \qquad (2.10)$$

$$\frac{\mathrm{d}Y_{i,j,k}}{\mathrm{d}t} = \theta_j \lambda_{i,j,k} X_{i,j,k} - \sigma_j Y_{i,j,k}, \qquad (2.11)$$

$$\frac{\mathrm{d}Z_{i,j,k}}{\mathrm{d}t} = (1-\theta_j)\lambda_{i,j,k} X_{i,j,k} - \kappa_j Z_{i,j,k}, \qquad (2.12)$$

where $\lambda_{i,j,k}$ is the force of infection experienced by members of subpopulation i of gender j and sexual AC k. The term for force of infection in the MP model has two key components,

$$I_{i,j} = \frac{\sum_k(c_{3-j,k}(Y_{i,3-j,k}+Z_{i,3-j,k}))}{\sum_k(c_{3-j,k}(X_{i,3-j,k}+Y_{i,3-j,k}+Z_{i,3-j,k}))}, \qquad (2.13)$$

the risk that a 'local sexual partnership' is with an infectious member of the opposite gender (i.e. $3-j$), and

$$E_j = \frac{\sum_{i,k}(c_{3-j,k}(Y_{i,3-j,k}+Z_{i,3-j,k}))}{\sum_{i,k}(c_{3-j,k}(X_{i,3-j,k}+Y_{i,3-j,k}+Z_{i,3-j,k}))}, \qquad (2.14)$$

the risk that a 'global sexual partnership' is with an infectious member of the opposite gender. The force of infection can thus be written as

$$\lambda_{i,j,k} = \beta_j c_{j,k}(\pi I_{i,j} + (1-\pi)E_j). \qquad (2.15)$$

The reproductive number can also be computed for the MP model but, in addition, the average number of infections caused by the typical infectious case through

'local sexual partnerships' in each subpopulation, R_i, can also be estimated. We shall first describe the expression for R_i, since it is very similar to the expression for R_0 in the AC model:

$$R_i^{\mathrm{M}} = \frac{\beta_2 d_1 \pi}{N_i} \begin{pmatrix} c_{1,1} D_{i,2,1} c_{2,1} & c_{1,1} D_{i,2,2} c_{2,2} \\ c_{1,2} D_{i,2,1} c_{2,1} & c_{1,2} D_{i,2,2} c_{2,2} \end{pmatrix}, \quad (2.16a)$$

$$R_i^{\mathrm{F}} = \frac{\beta_1 d_2 \pi}{N_i} \begin{pmatrix} c_{2,1} D_{i,1,1} c_{1,1} & c_{2,1} D_{i,1,2} c_{1,2} \\ c_{2,2} D_{i,1,1} c_{1,1} & c_{2,2} D_{i,1,2} c_{1,2} \end{pmatrix}, \quad (2.16b)$$

where $D_{i,j,k}$ is the number of individuals in subpopulation i of gender j and AC k, and N_i is the number of partnerships within each subpopulation, where

$$N_i = D_{i,1,1} c_{1,1} + D_{i,1,2} c_{1,2}$$
$$= D_{i,2,1} c_{2,1} + D_{i,2,2} c_{2,2}. \quad (2.17)$$

R_i is likewise the geometric mean of R_i^{M} and R_i^{F},

$$R_i = \sqrt{R_i^{\mathrm{M}} R_i^{\mathrm{F}}}. \quad (2.18)$$

The mathematical expressions for R^{M} and R^{F} for the whole MP model are more complex. For a MP model with 100 subpopulations,

2.3. Model parameters

To illustrate how the model performs, we used parameters (table 1) that describe gonococcal infection and sexual behaviour in the UK. For the former, we used the same per partnership transmission probabilities, and durations of infectiousness, as proposed by Garnett *et al.* (1999). However, because clinic-based observations overestimate the proportion of new infections that are symptomatic and care seeking (Farley *et al.* 2003), we used values from a study that estimated this parameter from the combined results of clinic and community-based case-finding methods (Farley *et al.* 2003). The total number of incident cases (T) was based on the modelled incidence of symptomatic care-seeking infections, divided by an adjustment factor since care-seeking infections account for only a fraction (α_j) of all clinic patients seen (Jungmann *et al.* 2004), i.e. $T = \sum_{i,j,k} ((\theta_j \lambda_{i,j,k} X_{i,j,k})/\alpha_j)$, and expressed as rates per 100 000 persons per year.

Data for sexual behaviour were derived from the National Survey of Sexual Attitudes and Lifestyles (NATSAL), a cross-sectional population-based survey of individuals aged 16–44 in the UK in the year 2000 (Johnson *et al.* 2001). Individuals who reported

$$R^{\mathrm{M}} = \beta_2 d_1 \begin{pmatrix} c_{1,1} D_{1,2,1} c_{2,1} \left(\frac{\pi}{N_1} + \frac{1-\pi}{N} \right) & c_{1,1} D_{1,2,2} c_{2,2} \left(\frac{\pi}{N_1} + \frac{1-\pi}{N} \right) & \cdots & c_{1,1} D_{100,2,2} c_{2,2} \left(\frac{1-\pi}{N} \right) \\ c_{1,2} D_{1,2,1} c_{2,1} \left(\frac{\pi}{N_1} + \frac{1-\pi}{N} \right) & c_{1,2} D_{1,2,2} c_{2,2} \left(\frac{\pi}{N_1} + \frac{1-\pi}{N} \right) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ c_{1,2} D_{1,2,1} c_{2,1} \left(\frac{1-\pi}{N} \right) & \cdots & \cdots & c_{1,2} D_{100,2,2} c_{2,2} \left(\frac{\pi}{N_{100}} + \frac{1-\pi}{N} \right) \end{pmatrix},$$
$$(2.19a)$$

$$R^{\mathrm{F}} = \beta_1 d_2 \begin{pmatrix} c_{2,1} D_{1,1,1} c_{1,1} \left(\frac{\pi}{N_1} + \frac{1-\pi}{N} \right) & c_{2,1} D_{1,1,2} c_{1,2} \left(\frac{\pi}{N_1} + \frac{1-\pi}{N} \right) & \cdots & c_{2,1} D_{100,1,2} c_{1,2} \left(\frac{1-\pi}{N} \right) \\ c_{2,2} D_{1,1,1} c_{1,1} \left(\frac{\pi}{N_1} + \frac{1-\pi}{N} \right) & c_{2,2} D_{1,1,2} c_{1,2} \left(\frac{\pi}{N_1} + \frac{1-\pi}{N} \right) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ c_{2,2} D_{1,1,1} c_{1,1} \left(\frac{1-\pi}{N} \right) & \cdots & \cdots & c_{2,2} D_{100,1,2} c_{1,2} \left(\frac{\pi}{N_{100}} + \frac{1-\pi}{N} \right) \end{pmatrix}.$$
$$(2.19b)$$

In the above equations, N is the total number of partnerships in all the subpopulations, as given by

$$N = \sum_{1}^{100} N_i. \quad (2.20)$$

R_0 is again computed as the geometric mean of R^{M} and R^{F} using next-generation operators.

partners of the same gender in the past 1 year were excluded, as were invalid responses, and weights for the core sample in NATSAL were factored in for all calculations. We used only observations from respondents living in Greater London as much of the other data referred to in this work were also from studies based in London (Low *et al.* 1997; Hickman *et al.* 1999;

Table 1. Model parameters.

characteristics of gonococcal infection	symbol	men	women	source of data
probability of acquiring the disease, per partnership	β_j	0.6	0.8	
duration of infectiousness, symptomatic care-seeking infections	$1/\sigma_j$	13 days	20 days	Garnett *et al.* (1999)
duration of infectiousness, non-care-seeking infections	$1/\kappa_j$	185 days	185 days	
proportion of new infections are symptomatic and care-seeking	θ_j	0.59	0.36	Farley *et al.* (2003)
proportion of clinical infections seek care from symptoms	α_j	0.89	0.53	Jungmann *et al.* (2004)

mixing between subpopulations	symbol	men	women	source of data
proportion of partnerships are in the same subpopulation, %	π	70% (50–90%)		(see text)

sexual activity, by no. of new partners in past year		men			women			source of data
		0–4	5+	all	0–4	5+	all	
proportion in activity state	$P_{j,k}$	96.35%	3.65%	100%	98.78%	1.22%	100%	NATSAL[a]
no. of new partners in past year from NATSAL, mean	$b_{j,k}$	0.58	10.16	0.93	0.31	6.15	0.38	
derived no. of new partners in past year used in model	$c_{j,k}$	0.45	6.11	0.66	0.44	18.23	0.66	(see text)

[a] Based on NATSAL sample for London (weighted $N=739$ for men and 728 for women).

Low 2002; Turner *et al.* 2004; Choudhury *et al.* 2006; Risley *et al.* 2007).

To obtain a guide as to how frequently sexual partnerships tend to occur within subpopulations (i.e. to obtain a sensible range of values for π), we looked at data on partnership choices pertaining to age, area of residence and ethnicity (table 1). NATSAL data reveal that, close to 90 per cent of first sexual partners, but only approximately 75 per cent of most recent partners, were of about the same age (± 5 years) as the respondent; further analyses suggested this was an age-related phenomenon, since the most recent partner for more than 90 per cent of younger persons (age less than 20) was of a similar age. NATSAL suggested that approximately 70 per cent of most recent partners were from the same town as the respondent, but this proportion varied when the data were sub-analysed by age and ethnicity. A study on ethnic related mixing also found that most individuals chose partners from the same ethnic group (Turner *et al.* 2004), while another study based on investigation of case-contact pairs of gonorrhoea in Colorado Springs in the USA also suggested that infected individuals had a tendency to select partners from the same ethnic group and geographical zones (as defined in that study) more than from other socio-geographic groups (Rothenberg & Potterat 1988). We decided we would explore the range of values from 50 to 90 per cent, with a base value of 70 per cent.

Table 1 also gives the proportion of men and women reporting 0–4 and 5+ new partners in the past 1 year, and the observed partner change rates in NATSAL

($b_{j,k}$). As has been noted by others (Johnson *et al.* 2001), the average number of new partners reported at a population level by men exceeded that for women. We shall explain our choice of using five or more new partners as our cut-off point, and our strategy for 'balancing' the number of new partnerships from men and women of each AC in subsequent sections.

2.4. Balancing male and female partnerships

The number of partnerships formed by one gender should equal that formed by the opposite gender at the population level. Because we intended to compare our MP model against the AC model, an additional constraint, which has factored into versions of the AC model of STI transmission (Garnett *et al.* 1999), is that the number of partnerships provided by men and women from the same AC, k, must also be balanced, i.e.

$$c_{1,k}P_{1,k} = c_{2,k}P_{2,k}. \tag{2.21}$$

We decided to use the average reported number of partnerships by both the men and women from each AC ($b_{j,k}$) to compute $c_{j,k}$,

$$c_{1,k}P_{1,k} = c_{2,k}P_{2,k} = (b_{1,k}P_{1,k} + b_{2,k}P_{2,k})/2. \tag{2.22}$$

The result is that partner change rates for the men are adjusted downwards, and those for the women adjusted upwards (table 1).

2.5. Generating a distribution of high-activity individuals across subpopulations, and model initialization and implementation

The Lorenz curve has been used to describe the asymmetric distribution of various social phenomena such as wealth across a population, and is calculated as follows:

$$L(F) = 1 - (1-F)^{1-\phi}, \qquad (2.23)$$

where F is a proportion between 0 and 1, and ϕ is a coefficient that determines how asymmetric the distribution is. When $\phi = 0$ we obtain a perfectly even distribution, whereas when $\phi = 1$ we obtain a perfectly uneven distribution. In our case, we used this function to describe how high-activity individuals might be distributed into the different subpopulations. The above equation was used to define 100 discrete levels of high-activity state individuals; when ranked in ascending order in terms of the density of high-activity state individuals, the cumulative proportion of all high-activity state individuals up to the rank, r, would then be

$$L_r = 1 - \left(1 - \frac{r}{100}\right)^{1-\phi}. \qquad (2.24)$$

Let U_r be the proportion of high-activity individuals within the level of rank r. U_r can be derived by subtracting the cumulative proportion up to the previous rank from the cumulative proportion for the current rank, i.e.

$$U_r = L_r - L_{r-1}. \qquad (2.25)$$

Therefore, if we assume that all the 100 subpopulations are the same size, then the number of high-activity men ($D_{i,1,2}$) in each subpopulation would be

$$D_{i,1,2} = U_i \times P_{1,2} \times S/2, \qquad (2.26)$$

where S is the total number of individuals in each subpopulation, of which half are assumed to be men. The remaining men in the subpopulation must be in the low-AC, hence

$$D_{i,1,1} = S/2 - D_{i,1,2}. \qquad (2.27)$$

If we assume that the supply of partnerships by men and women in each subpopulation is balanced, and the subpopulation size has been fixed, then

$$D_{i,1,1}c_{1,1} + D_{i,1,2}c_{1,2} = D_{i,2,1}c_{2,1} + D_{i,2,2}c_{2,2}, \qquad (2.28)$$

$$D_{i,2,1} + D_{i,2,2} = S/2. \qquad (2.29)$$

Since we set $c_{i,k}$ to be the same across all subpopulations, equations (2.28) and (2.29) can be solved to give $D_{2,i,1}$ and $D_{2,i,2}$, which are the number of women in the low and high-ACs for each subpopulation. The result is that, within each subpopulation, the total number of partnerships by men and women is balanced, but the contribution by men and women of the same AC can differ; i.e. men of one AC can contribute more or fewer partnerships than women of the same AC, and it is thus not possible to have fully assortative mixing within subpopulations. For simplicity, mixing within subpopulations was therefore assumed to be completely random.

Figure 2 illustrates the distribution of high-activity individuals, when using $\phi = 0.40$. Approximately 20 per cent of subpopulations would have a higher proportion of high-activity individuals than the population average, with the remaining subpopulations having slightly less than the population average.

We modelled a total population of 100 000 men and 100 000 women distributed into 100 subpopulations, with the above method for distributing high-activity individuals. We initialized the model by seeding the subpopulation with the most number of high-activity individuals with one non-care-seeking infection among the high-activity men; unless otherwise specified, results are based on the modelled incidence 50 years after initialization, a point where model outputs had reached equilibrium values.

2.6. Use of the Gini coefficient as a summary measure for the distribution of infections across subpopulations

Lorenz curves have been used to visualize and compare the distribution of STIs across geographically defined subpopulations (Elliott *et al.* 2002; Kerani *et al.* 2005; Monteiro *et al.* 2005). In such studies, the Lorenz curves plot the cumulative proportion of incident infections (y-axis) against that of the population (x-axis), after ranking population subunits (e.g. census wards) in order of their incidence rates. It is also possible to derive a summary measure for the degree to which cases are asymmetrically distributed between subpopulations, and this has been done by the above-mentioned studies for gonorrhoea, chlamydia and other STIs. This is computed by first defining a straight diagonal line ($x = y$), referred to as the line of equality. The extent to which Lorenz plots deviate from the line of equality (i.e. the result when cases are distributed equally across all subpopulations) is then quantified in terms of the Gini coefficient; this is twice the difference between the area under the curve for the modelled cases and the line of equality. For our MP model, the area in each interval beneath the curve for modelled cases approximates a trapezium and so, if we let G be the Gini coefficient, and H_i be the cumulative proportion of cases produced by each subpopulation ranked in descending order by incidence rates, then

$$G = 2 \times \left(\sum_{i=1}^{n} \left[\left(\frac{H_i - H_{i-1}}{2} + H_{i-1} \right) / n \right] - 0.5 \right). \qquad (2.30)$$

The Gini coefficients from our model were then compared with empirical estimates from a study on the demographic and geospatial incidence of gonorrhoea and three other STIs in Leeds (Monteiro *et al.* 2005), which, like London, is also an urban city area in the UK.

3. RESULTS

3.1. Persistence of infection in a single subpopulation

We begin by considering the infection dynamics within a single subpopulation. Let the proportion of high-activity individuals be $P_{j,2}$ with the remainder being low-activity individuals (i.e. $P_{j,1} = 1 - P_{j,2}$). For a
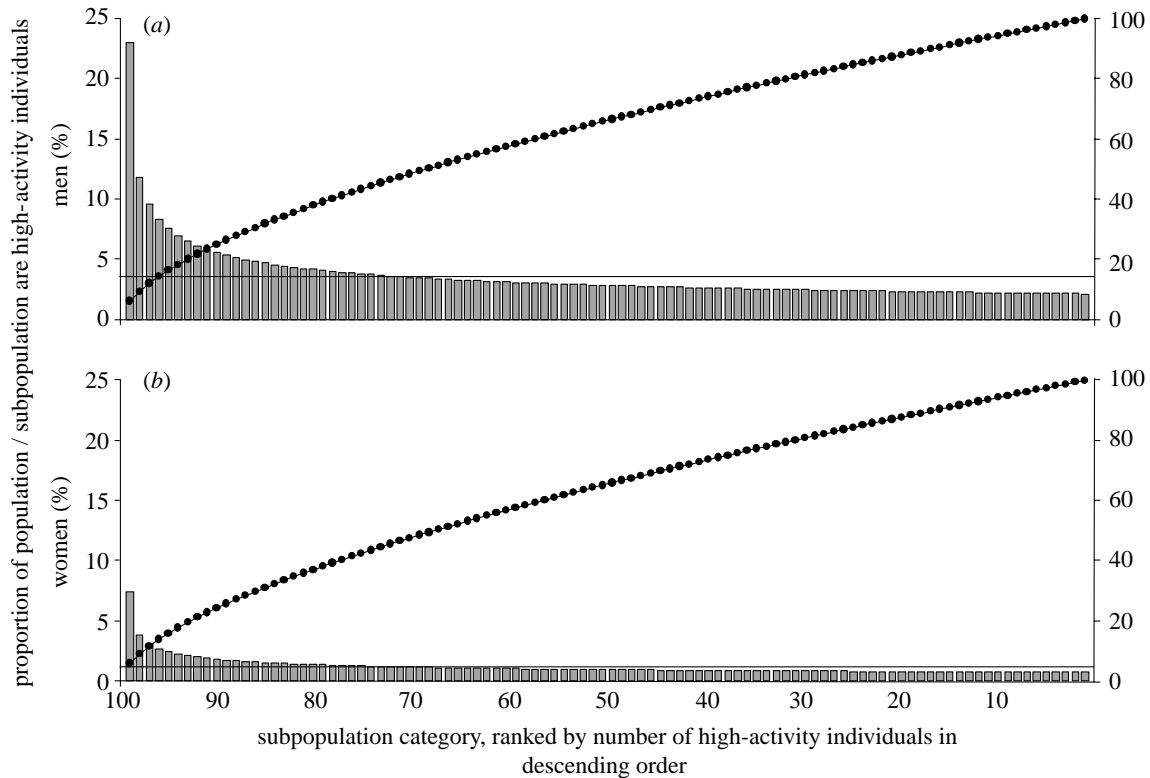
Figure 2. (*a*,*b*) Proportion of subpopulations of high-activity individuals (grey bars), proportion of general population of high-activity individuals (solid horizontal lines) and cumulative proportion of high-activity individuals (filled circles) when using $\phi = 0.40$. Across all subpopulations, there are fewer women than men in the high-AC, owing to the corresponding gender differences in the proportion of high-activity individuals observed in NATSAL (table 1). Approximately 20% of subpopulations would have a higher proportion of high-activity individuals than the population average, with the remaining subpopulations having slightly less than the population average. Different distributions can be generated, using different values for the distribution parameter, ϕ.

disease such as gonorrhoea, which in the UK has a reasonably low prevalence at the population level (Lavelle *et al.* 2006), the risk of catching infection by forming a partnership with an infectious individual from the population is low. As $E_j \rightarrow 0$, equation (2.15) can be reduced to

$$\lambda_{i,j,k} = \beta_j c_{j,k} \pi I_{i,j}. \tag{3.1}$$

By substituting equation (2.15) with (3.1), we modelled a single subpopulation while assuming all sexual partnerships with the rest of the population are 'wasted', and do not result in any infections (i.e. $E_j = 0$). Modelling a single subpopulation in this manner allows us to illustrate the relationship between π, the degree of within-subpopulation mixing, $P_{j,2}$, the proportion of high-activity individuals in the subpopulation being modelled, and the threshold value of $c_{j,2}$ needed for an infection to persist when the prevalence of infectious contacts from global sexual partnerships is negligibly low. To do so in the simplest possible manner, we assumed here that men and women behaved similarly ($P_{j,k} = P_{1,k} = P_{2,k}$, $c_{j,k} = c_{1,k} = c_{2,k}$), and fixed the partner change rates for low-activity individuals at a suitably low number ($c_{1,1} = c_{2,1} = 0.5$ per year).

When the model is run with gonorrhoea-like parameters, we see from figure 3*a* that, regardless of the value of π assumed, the threshold value of $c_{j,2}$ for persistence is very high when the proportion of high-activity individuals is small (less than 1%); the threshold partner change rate decreases rapidly when the proportion of high-activity individuals is increased

slightly, and then reaches a plateau when the proportion of high-activity individuals is increased further. NATSAL data (figure 3*b*) show that extreme partner change rates (>10 or more per year) are observed only with a small proportion of men (less or equal to 2%), and the highest partner change rate reported for women was 10 per year. We decided to use five or more partners as the cut-off point for classifying high-activity individuals, as this would give a partner change rate in high-activity individuals that allows persistence of infection in subpopulations with sufficiently high proportions of such individuals.

3.2. Calibrating the MP model

Based on our chosen parameter set, the proportion of high-activity individuals in the population would be too low to sustain gonococcal transmission. However, the crux of our MP model is the assumption that high-activity individuals are distributed asymmetrically across the population. Subject to the degree of within subpopulation mixing assumed (π), the distribution of high-activity individuals (ϕ), if asymmetric, could result in some subpopulations having sufficient concentrations of high-activity individuals to sustain transmission, even without the need for extreme partner change rates. This section thus explores how the two key model parameters, π and ϕ, interact in the MP model, and attempts to arrive at some sensible estimates for these two parameters.
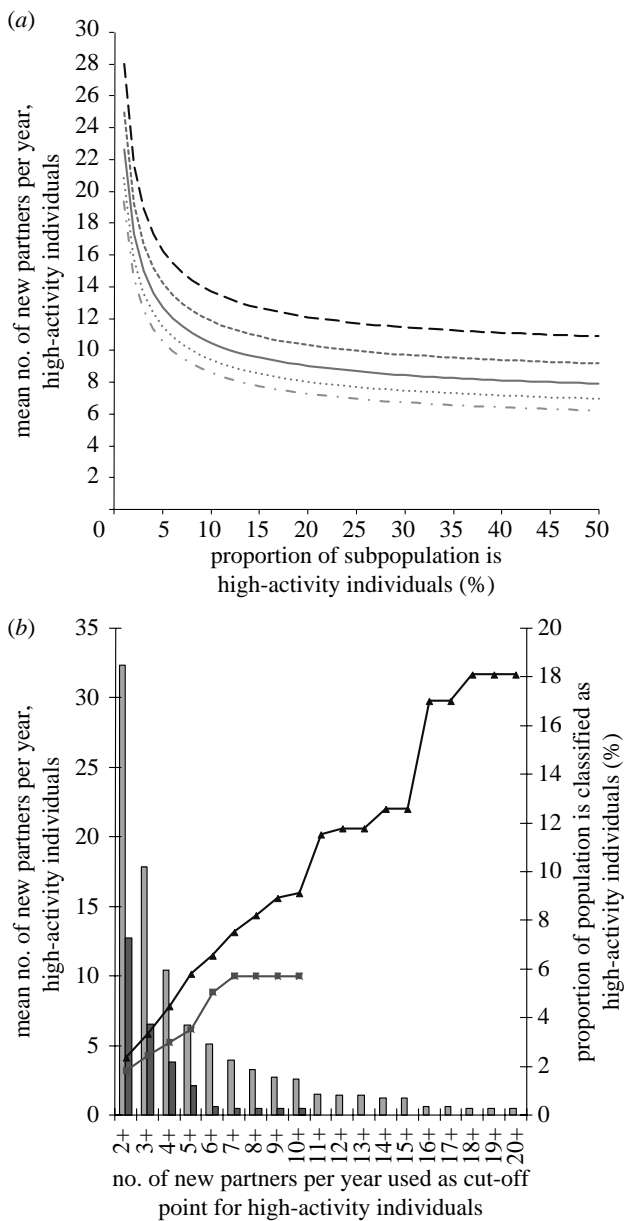
Figure 3. (*a*) Threshold partner change rate for high-activity individuals, for subpopulations with different proportions of high-activity individuals, and different proportions of sexual partnerships are exclusively from within the same subpopulation (*π*) (long-dashed curve, 0.5; short-dashed curve, 0.6; solid curve, 0.7; dotted curve, 0.8, dashed dotted curve, 0.9), assuming that all partnerships with the rest of the population are wasted (i.e. $E_j = 0$). Low-activity individuals are assumed to have a partner change rate of 0.5 per year. Subpopulations with a higher proportion of high-activity individuals, or where a greater proportion of sexual partnerships are with other individuals in the same subpopulation, would require a lower threshold partner change rate for a self-sustaining epidemic of gonorrhoea. (*b*) Mean number of new partners in high-AC individuals when assuming different cut-off points, based on NATSAL data for respondents from London. Extreme partner change rates (more than 10 per year) are observed only with a small proportion of men (less than 2%), and the highest partner change rate reported for women was 10 per year (light grey bars, percentage of men; dark grey bars, percentage of women; triangles, partner change, men; squares, partner change, women).

Figure 4*a* shows runs of the model while varying the degree of asymmetric distribution (*φ* from 0.3 to 0.6) and fixing the amount of within-subpopulation mixing at *π* = 0.7. More asymmetric distributions (higher values of *φ*) result in a higher overall incidence and reach the equilibrium incidence faster. Figure 4*b* presents the modelled incidence 50 years after seeding the population for different distributions of high-activity individuals (*φ*) and different levels of within-subpopulation mixing (*π*). The incidence rate for heterosexually acquired gonorrhoea in London, in the age groups 16–44, was approximately 200 cases per 100 000 population for the year 2001, i.e. around the time NATSAL data were collected (Health Protection Agency 2005); we see that multiple combinations of *φ* and *π* could reproduce the observed incidence. However, different values of *π* give rather different results for the distribution of modelled cases across subpopulations (figure 4*c*). Assuming higher levels of within-subpopulation mixing resulted in a more unequal distribution of cases, in spite of fixing the intensity of transmission at 200 cases per 100 000 population per year (by using the appropriate values of *φ* as obtained from figure 4*b*).

The Gini coefficient for gonorrhoea has been measured to be approximately 0.49 (Monteiro *et al.* 2005), suggesting that values of *π* of approximately 0.7 are the most appropriate given our model assumptions. These estimates corroborate the estimates on the proportion of sexual partnerships formed between members of identifiable subgroups in the behavioural data. Figure 4*d* further explores the relationship between within-subpopulation mixing, the distribution of high-activity individuals, and the resultant incidence and distribution of cases across subpopulations as measured through the Gini coefficient. With our chosen parameter set, it would appear that the degree of within-subpopulation mixing (*π*) is the more important determinant of the distribution of cases (as reflected by the Gini coefficient), even without the constraints of having to fit the model to the appropriate incidence (as given in the contour lines). For instance, values of *π* below 0.5 are unable to give Gini coefficient values consistent with what was measured for gonorrhoea, unless extreme values of *φ* are assumed—such values in turn would imply a subpopulation composed almost entirely of high-activity men. As such, Gini coefficients derived from the observed distribution of STIs may be one way of supplementing estimates on the degree of within-subpopulation mixing obtained from behavioural studies.

In all subsequent analyses, we assumed a value of *π* = 0.7 and calibrated the model to give an incidence rate of 200 cases per 100 000 population by setting *φ* to 0.4032. These parameter values gave a Gini coefficient of 0.47, which is close to the result obtained by Monteiro *et al.* (2005).

3.3. Comparison with AC model of STI transmission

We implemented the AC model using parameters identical to those used in our MP model. In a manner similar to adjusting *φ* in the MP model, *ε* can be
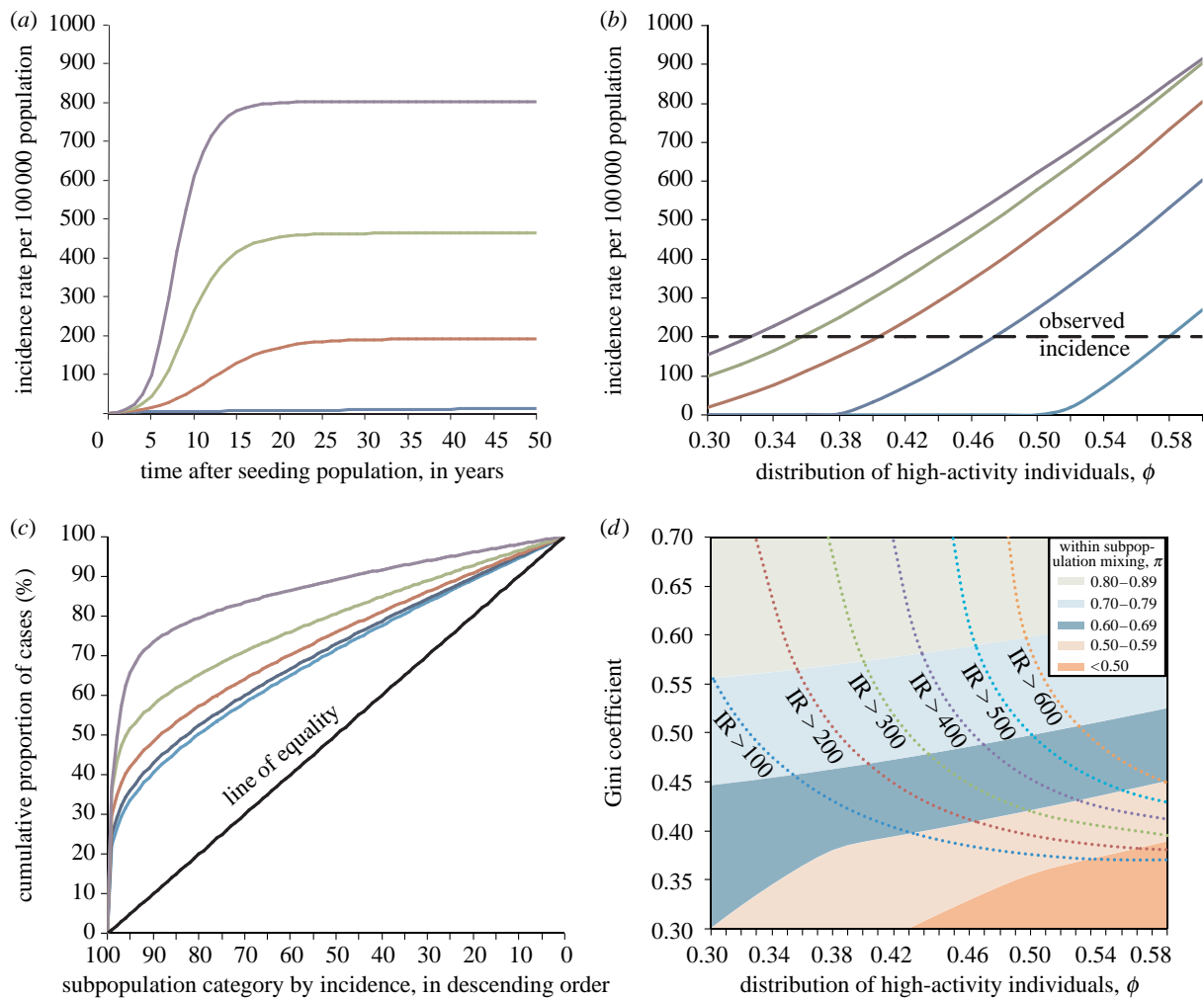
Figure 4. (*a*) Distribution of high-activity individuals (ϕ) and endemic incidence for $\pi = 0.7$, where π is the proportion of sexual partnerships formed exclusively with those from within the same subpopulation. Lines denote different distributions of high-activity individuals. More asymmetric distributions (higher values of ϕ; blue curve, $\phi = 0.3$; red curve, $\phi = 0.4$; green curve, $\phi = 0.5$, violet curve, $\phi = 0.6$) cause a higher incidence at equilibrium, and reach this equilibrium faster. (*b*) Distribution of high-activity individuals (ϕ) and incidence at 50 years for different levels of within-subpopulation partnerships (π, depicted by coloured lines; light blue curve, $\pi = 0.5$; dark blue curve, $\pi = 0.6$; red curve, $\pi = 0.7$; green curve, $\pi = 0.8$; violet curve $\pi = 0.9$). Dashed line shows approximate incidence observed for London, 2001. For plausible ranges of π, the observed incidence can be reproduced by an appropriate value for the distribution of high-activity individuals. (*c*) Lorenz curves giving cumulative distribution of cases with different levels of within-subpopulation partnerships (π) (light blue curve, $\pi = 0.5$ (Gini = 0.38); dark blue curve, $\pi = 0.6$ (Gini = 0.41); red curve, $\pi = 0.7$ (Gini = 0.47); green curve, $\pi = 0.8$ (Gini = 0.56); violet curve $\pi = 0.9$ (Gini = 0.73); black line, line of equality), with incidence fixed at 200 cases per 100 000 population, using the appropriate values of ϕ from (*b*). When subpopulations are sorted by the number of cases they produce, it becomes clear that a minority of subpopulations contribute to more cases. Deviation from the line of equality (i.e. when cases are distributed equally across all subpopulations) can be quantified in terms of the Gini coefficient (see text). Assuming that more sexual partnerships occur exclusively within the same subpopulation (π) would result in more unequal distribution of cases. (*d*) Distribution of high-activity individuals (ϕ) and levels of within-subpopulation partnerships (π) on incidence and Gini coefficient. Areas shaded by different colours denote the boundaries observed when assuming different levels of within-subpopulation partnerships (π) for different distributions of high-activity individuals (ϕ). Each pair of values for within-subpopulation partnerships (π) and distribution of high-activity individuals (ϕ) result in a different distribution of cases and hence a different Gini coefficient (*y*-axis), as well as a different incidence rate; the boundaries between different incidence rates per 100 000 population (IR) are given by the contour lines. To obtain the Gini coefficient of 0.49 observed for gonorrhoea while constraining the incidence to approximately 200 per 100 000 per year, π should lie between 0.6 and 0.8, with a corresponding value of ϕ between 0.36 and 0.48.

adjusted to set the incidence for the AC model; as has been reported by Garnett *et al.* (1999), we found that near-random mixing ($\varepsilon = 0.2158$) was required for the model to produce realistic incidence rates.

Figure 5 illustrates several key differences in the dynamics of the two models. On seeding the two models with an infectious case (figure 5*a*), the MP model reaches its equilibrium incidence within approximately

50 years, but it takes almost 300 years for this to happen in the AC model. We also demonstrate how making small changes in parameter estimates changes the modelled incidence. Firstly, waiting times for accessing care for STIs in the UK were noted to have increased (Djuretic *et al.* 2001; Foley *et al.* 2001; Griffiths *et al.* 2004) in the last decade. We modelled a 5-day increase for the duration of care-seeking
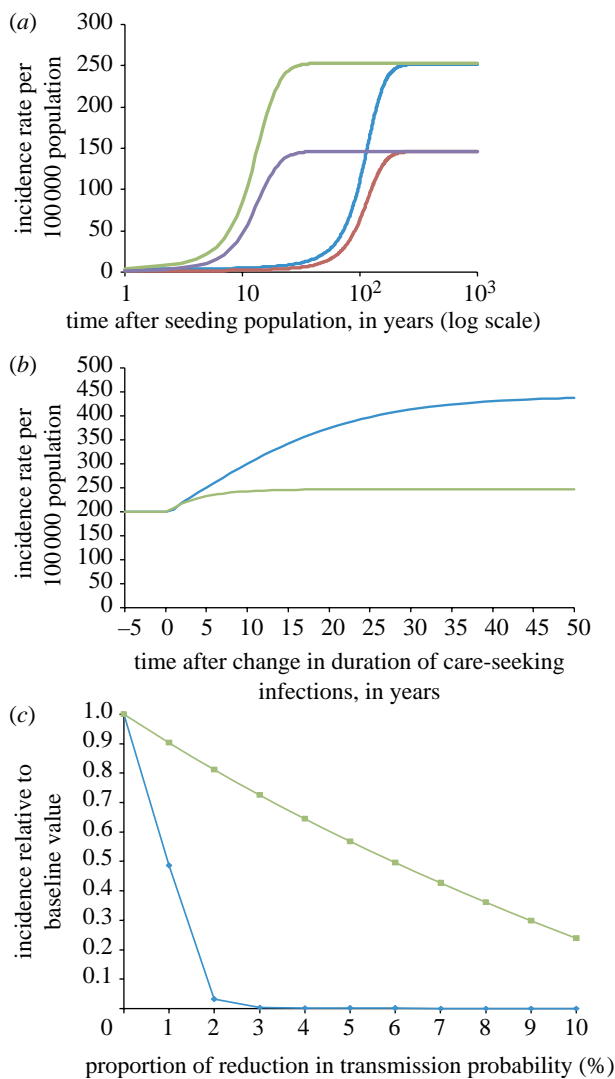
Figure 5. (*a*) Modelled change in incidence rates over time after seeding population. The mixing parameter for the AC model (blue curve, men; red curve, women) is set to $\varepsilon = 0.2158$, to give an incidence rate of 200 cases per 100 000; the MP model (green curve, men; violet curve, women) uses $\pi = 0.7$ and $\phi = 0.4032$. The MP model reaches its equilibrium incidence within approximately 50 years, but it takes more than 200 years for this to happen in the AC model. (*b*) Modelled change in incidence rates over time following a 5-day increase in the duration of care-seeking infections; other parameters as for (*a*). The MP model (green curve) predicts that incidence will increase by approximately 20% and stabilize at a new equilibrium in approximately 10 years, whereas the AC model (blue curve) suggests that incidence would more than double, and take approximately 50 years to reach the new equilibrium. (*c*) Change in equilibrium incidence with up to 10% reduction in transmission probability, with incidence expressed as a proportion relative to the baseline incidence. Equilibrium incidence was approximated by running the models until year-on-year change in incidence was less than 0.01%; other parameters are as for (*a*). The AC model (diamonds) predicts much larger reductions in incidence than for the MP model (squares).

infections, consistent with the variations in time needed to access care which have been observed (Griffiths *et al.* 2004). The MP model predicts that incidence would increase by 20 per cent and reach a new equilibrium in 5–10 years, but the AC model suggests that incidence

would continue to increase even after 50 years, and reach a new equilibrium, which is more than double the baseline incidence (figure 5*b*). Secondly, the NATSAL survey in 2000 revealed an increase in consistent condom use over the corresponding survey of sexual lifestyles in 1990 (Johnson *et al.* 2001). We therefore modelled how the equilibrium incidence would change with a decrease in per partnership transmission probability. Figure 5*c* shows that a 1–10 per cent reduction in transmission probability would result in a near proportionate decrease in incidence for the MP model; the AC model, however, predicts that even 1 per cent reductions in transmission probability would lead to a sharp decrease in incidence, and that transmission is no longer sustained with just a 3 per cent reduction in transmission probability.

3.4. Comparison of R_0 for AC and MP models

One way of explaining the difference in dynamics between the AC and MP models is to examine the values of R_0, the reproductive number, for the two models. Although the two models have been calibrated to produce the same incidence rates, the MP model actually has a higher R_0 than the AC model. The values of R_0 for the AC and MP models are given below.

In the AC model, for $\varepsilon = 0.2158$, which gave an incidence rate of approximately 200 infections per 100 000 population a year in the dynamic formulation of the AC model, $R_0 \approx 1.02$, suggesting that the infection would be at the threshold of persistence. In the MP model, for $\pi = 0.7$ and $\phi = 0.4032$, which also produces 200 infections per 100 000 population a year, $R_0 \approx 1.17$, which is reasonably higher than 1. Values of R_i can also be computed; these range from 0.39 in the subpopulation with the least high-activity individuals ($i = 1$) to 1.15 in the subpopulation with the most high-activity individuals ($i = 100$). We see that the highest value of R_i is less than the value of R_0 for the entire population, since the formula for R_i accounts only for infections caused from local sexual partnerships in that subpopulation, and not for the small contribution of infections from global sexual partnerships. We can also infer that, since the maximum value of R_i very nearly approaches the value of R_0 for the entire population, the transmission dynamics in the MP model is mainly influenced by this subpopulation.

Owing to the higher value of R_0 in the MP model, the infection propagates much faster than in the AC model once the infection has been introduced (figure 5*a*). The variation in R_i also means that parameter changes increasing transmission affect mainly the few subpopulations where R_i is nearly 1 or greater than 1. The 5-day increase in duration of care-seeking infections in figure 5*b* increases R_0 to 1.20 in the MP model, and 1.04 in the AC model. However, in the MP model, R_i remains below 1 in most subpopulations, so that we observe a much smaller change in incidence (figure 5*b*) than in the AC model where the increase in transmission occurs across the entire population. And finally, with small parameter changes that reduce
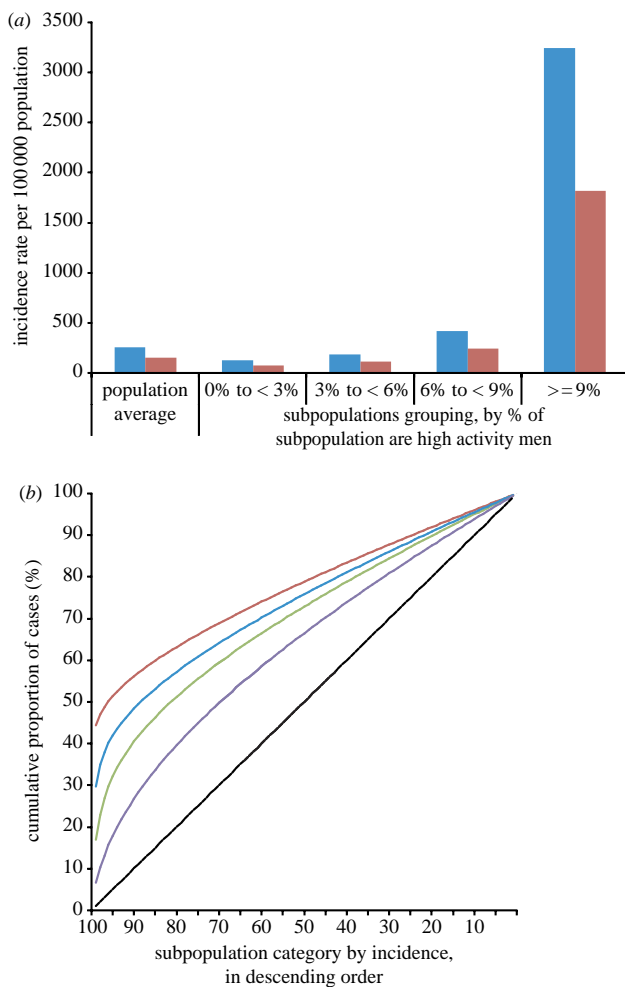
Figure 6. (*a*) Modelled incidence rate for subpopulations grouped by the proportion of high-activity men. Parameters are $\pi = 0.7$ and $\phi = 0.4032$. Subpopulations with a low proportion of high-activity men (less than 6%) have an incidence rate lower than the population average, whereas the subpopulations with the highest proportion of high-activity men (more than 9%) have an incidence rate more than 10 times the population average (blue bars, men; red bars, women). (*b*) Lorenz curves using the MP model for gonorrhoea-like parameters, lower and higher transmissibility, and chlamydia-like parameters. π and ϕ are as for (*a*). A less transmissible disease (15% reduction) would be more highly concentrated in subpopulations that have a higher proportion of high-activity individuals, whereas a more transmissible disease (15% increase) would be more evenly spread out across the subpopulation categories. Chlamydia-like parameters produce a pattern of infections that is spread out more evenly than gonorrhoea (red curve, $\downarrow\beta^{\mathrm{M}}$, β^{F} (Gini = 0.54); green curve, $\uparrow\beta^{\mathrm{M}}$, β^{F} (Gini = 0.40); blue curve, gonorrhoea (Gini = 0.47); violet curve, chlamydia (Gini = 0.26); black dashed line, line of equality).

transmission (figure 5*c*), R_0 remains above 1 in the MP model, but the already low value of R_0 is brought below 1 in the AC model. For example, for a 2 per cent reduction in the probability of transmission, $R_0 \approx 1.14$ in the MP model, but $R_0 \approx 1.00$ in the AC model; and while the MP model tolerates a 15 per cent reduction in transmission probability before reaching the threshold below which no self-sustaining transmission occurs, a 3 per cent reduction leads to $R_0 < 1$ in the AC model.

3.5. MP explanation for epidemiological observations about gonorrhoea and other STIs

In this final section of the results, we suggest how the MP model potentially explains several epidemiological observations about gonorrhoea and other STIs. Figure 6*a* presents the modelled incidence rate for subpopulations grouped by the proportion of high-activity men. Not surprisingly, modelled incidence rates in subpopulations where 9 per cent or more of men are in the high-AC are more than 10 times the population average, although we have let the high- and low-activity individuals in these subpopulations have the same partner change rates as those in other subpopulations. Notably, individuals in the low-AC in such subpopulations would have an elevated risk of infection, and this may explain the contextual risk that has been observed based on various sociodemographic factors such as age and ethnicity (Fenton *et al.* 2005), which probably influence subpopulation membership. Figure 6*b* revisits the Lorenz curves, where gonorrhoea is compared against an STI with a lower transmissibility, and one with higher transmissibility, as well as a parameter set depicting infections with *Chlamydia trachomatis*. All Lorenz curves and Gini coefficients were derived in the same way as was described previously, again using $\pi = 0.7$ and $\phi = 0.4032$, values found in earlier analysis to give a reasonable depiction for the incidence and distribution of gonorrhoea. An STI with similar infectious duration, but lower transmissibility, i.e. a lower R_0, would be less evenly distributed than gonorrhoea, and an STI with higher transmissibility would be more ubiquitous. To model chlamydia, we assumed that transmissibility was fairly similar (Lin *et al.* 1998), but that only 9 per cent and 24 per cent of incident male and female infections, respectively, would seek care. These estimates came from the same study from which we obtained the corresponding parameter for gonorrhoea (Farley *et al.* 2003) and all other parameters were left unchanged. Monteiro and colleagues had previously estimated a Gini coefficient of 0.26 for chlamydia (Monteiro *et al.* 2005). This tallies well with the result from the MP model, providing further support for our underlying assumptions about the distribution of risk behaviour in the population.

4. DISCUSSION AND CONCLUSION

Since the seminal work of Hethcote and Yorke more than 20 years ago, there have been different interpretations as to what constitutes the core group in STI transmission, both in epidemiological as well as in modelling studies (some of which were highlighted earlier in our paper). In some ways, our work provides insights similar to the original concept of the core group by Hethcote and Yorke, as the MP model also proposes that the dynamics of certain STIs is driven mainly by a few small groups of individuals, with the rest of the population contributing little to the maintenance of transmission. However, our work improves on the original model in several ways. For one, the model

presented potentially explains other key epidemiological features of gonorrhoea in the UK, such as the uneven sociodemographic (Lacey *et al.* 1997; Low *et al.* 1997) and geographical (Monteiro *et al.* 2005) distribution of gonococcal disease. By illustrating how individuals with low sexual risk activity can have a high contextual risk of infection due to their subpopulation membership and vice versa, the model also provides a framework for understanding why the risk for bacterial STIs remains associated with demographic (Fenton *et al.* 2005) and geographical factors (Hardick *et al.* 2003; Jennings *et al.* 2004) after adjusting for the sexual behaviour of the individual. Moreover, the MP model can produce realistic incidence rates for gonorrhoea, while reasonably representing the effects of small changes in parameters such as infectious duration and transmission probabilities, which are difficult to achieve with the AC model. In addition, the MP model can potentially be applied to other STIs. There have been attempts to generalize the concepts in the model by Hethcote and Yorke to offer explanations as to why the distribution of different STIs varies within the same population (Brunham & Plummer 1990), and our MP modification to the model by Hethcote and Yorke explains how this might be explained by a continuum of groups with different concentrations of risk activity. We believe that the above characteristics of our MP model constitute significant improvements over the 'AC' formulation of the Hethcote and Yorke model.

The issues with using the AC model for modelling gonorrhoea were actually identified by Garnett some years ago (Garnett *et al.* 1999; Garnett & Bowden 2000): 'the observed prevalences of infection for gonorrhoea and syphilis in industrialized countries suggest that the infections are near the threshold where they can persist. In such circumstances, the model fails to realistically represent interventions because elimination in the model is too easy.' Indeed, a replica of the AC model with our parameter set showed that this was indeed the case, with $R_0 \approx 1.02$. Note that we did assume lower proportions of care-seeking infections than did Garnett in his 1999 work, who modelled 95 per cent and 60 per cent of infections in men and women, respectively, as care-seeking. However he assumed a much higher maximum partner change rate of 40 new partners per year (Garnett *et al.* 1999). Had we used parameter estimates closer to these ($\theta_1 = 0.95$, $\theta_2 = 0.60$, $P_{1,2} = P_{2,2} = 0.0055$, $c_{1,2} = c_{2,2} = 40$, $c_{1,1} = c_{2,1} = 1.3$ to represent the average for other ACs), we could still have obtained an incidence of 200 per 100 000 cases per year (but at $\phi = 0.3737$), with the resultant model being just as robust to parameter changes as with our original parameter set. Garnett and Bowden recognized that the key failing of the AC model for low-prevalence situations was not in the parameter sets used and suggested in their work in 2000 that the answer to modelling a low-prevalence STI might lie in spatial stratification and sex-partner network structures. This is reasonably well captured by the MP framework we have proposed.

While MP depictions of STI transmission have been published before, those models were mainly for the purpose of modelling inter-regional spread of STIs

(Thomas & Smith 2000; Grassly *et al.* 2005). Our model provides a novel but effective way of modelling low-prevalence STIs while simultaneously explaining the asymmetric distribution of cases within a population. The critical feature of our model is the underlying MP structure with risk activity asymmetrically distributed across the subpopulations. This asymmetric distribution could possibly be the result of interactions between demographic and geographical factors, although our model does not explicitly identify these factors. Others have produced reasonably realistic models for gonorrhoea in the UK, by incorporating demographic factors such as age (White *et al.* 2005) and ethnic strata (Turner *et al.* 2004) into their models. Such work also implicitly acknowledged that there were subpopulations at higher risk which were potentially driving gonococcal transmission and our MP framework may appear to be simply a more generic way of viewing such underlying dynamics. However, we propose that there are critical differences between our model and models stratified by age and ethnicity. Firstly, without the MP framework, all individuals in the high-activity state within a stratum have an equal chance of mixing with each other, and would contribute equally to gonococcal transmission within their respective stratum. With the MP framework, however, only a small proportion of individuals in the high-activity state (those in subpopulations with self-sustaining transmission) are critical to spreading the infection, with the majority of high-activity individuals being protected from infection by being members of subpopulations where self-sustaining transmission does not occur. Secondly, when calibrated to give the low incidence rates typical for gonorrhoea in developed Western countries, stratified models may be dependent on one particular demographic stratum, such as the age group with the highest risk activity, to maintain transmission. This is not borne out in molecular epidemiologic data, which instead suggests that distinct strains circulate in multiple sexual networks with very different demographic features (Choudhury *et al.* 2006).

The findings from the MP model do, however, find parallels in prior work on individual-based and network models of STI transmission. For instance, work on individual-based models suggests that minor epidemics of gonorrhoea are sustained within small sections of the population, and simulating larger populations increases the likelihood of gonorrhoea finding a niche (Ghani *et al.* 1997). These pockets of transmission essentially occur within larger 'network components', with the infection being more likely to become established in populations with a higher component size. In the case of the MP model, the subpopulations with a higher density of high-activity individuals correspond somewhat to the network components that provide the niche for gonococcal transmission. In addition, the contextual risk of infection as described in the MP model (figure 6a) has a parallel in the finding from network models on the importance of local network structures, network position, and the risks of the sexual partners of an individual in determining how at risk an individual is of acquiring an infection (Ghani & Garnett 2000). It could be argued that individual-based and network

models demonstrate some of these concepts with much finer resolution, and are able to model factors critical to STI transmission such as the contribution of concurrent partnerships to transmission (Kretzschmar 2000) and the potential impact of partner notification as an intervention (Kretzschmar 2000; Turner *et al.* 2006). These are factors that cannot be modelled in the MP framework that has been presented. On the other hand, some have commented on the difficulties in using individual-based and network models to develop STI prevention strategies, and given several reasons as to why this might be so. Kretzschmar, in her paper on sexual network structure (Kretzschmar 2000), sums it up as follows: 'First, we simply do not yet have a good understanding of the relationship between network structure and disease transmission. Second, targeting larger network structures, such as connected components, is more difficult than targeting individuals and necessitates a deeper knowledge of global network structure.' In addition, Garnett has pointed out that individual-based models carry 'significant costs in both identifying parameter values and interpreting which measures are significant in generating results' (Garnett 2002). We propose that the MP approach complements individual-based and network models in contributing to our understanding of STI transmission and the potential impact of control measures in several ways. Firstly, the MP model suggests that the network components critical to transmission are possibly the result of clusters of connected individuals arising from shared sociogeographic characteristics, and that any attempt to build realistic individual-based and network models should thus consider whether the population modelled has some intrinsic MP structure along sociogeographic lines. Secondly, the MP model suggests that individual-based modelling approaches could be used in an attempt to identify the sociogeographic characteristics that contribute most to the formation of such clusters and components. Identifying the sociogeographic characteristics most pertinent to transmission may in turn yield target groups for future research and preventive interventions. Finally, parameter estimates derived from fitting MP models to rates of STIs by geographical or sociodemographic groups in conjunction with pertinent data on sexual behaviour could help to corroborate information on behavioural parameters such as the extent of mixing within and between different population subgroups. Such parameter estimates could then be used to guide the construction of more relevant individual-based and network models.

As with other modelling work, our model has some limitations. Firstly, the parameters describing the characteristics of gonococcal infection remain uncertain, and different assumptions about these parameters would affect the incidence and distribution of cases produced in this or any other model of gonococcal transmission. In addition, studies in different settings have estimated different values for the Gini coefficient. One Canadian study calculated a value of 0.66 for Winnipeg, Manitoba (Elliott *et al.* 2002), and one USA-based study calculated a value of 0.57 for King County, Washington (Kerani *et al.* 2005). The different estimates obtained suggest that estimates of the Gini

coefficient for an STI could be dependent on the methodology used, the context of the data, or both factors, and questions could therefore be raised about the accuracy and the validity of calibrating our model outputs against the Gini coefficient published by Monteiro and colleagues, which was based on data from Leeds from 1994 to 1995 (Monteiro *et al.* 2005). However, as alluded to earlier in our discussion, changes in either the parameter values or the estimate of the Gini coefficient that our model was calibrated to would not change the underlying dynamics assumed in the MP framework, and the result would simply be that different estimates are derived for the within-subpopulation mixing and distribution coefficients (π and ϕ, respectively). Another limitation is that our model considered only two sexual ACs, and fixed the partner change rate for each AC. This could be contentious, since there is a wide variation in partner change rates in the population, but we would suggest the key concept represented here is that transmission is being driven by subpopulations with a higher concentration of individuals with a combination of risk factors that facilitate transmission (most likely high partner change rates, but possibly other factors such as delayed care-seeking), rather than being driven by a few individuals with extreme levels of risk as is depicted in the AC models. In fact, we believe that attributing maintenance of transmission to individuals of extremely high partner change rates, such as implied by work proposing that scale-free networks drive STI transmission (Schneeberger *et al.* 2004), would be inconsistent with empirical data (figure 3b). In addition, models that use per sex act rather than per partnership transmission probabilities suggest that individuals with high partner change rates but single sexual encounters with many partners may be less potent to STI transmission than individuals with lower partner change rates but who have several sexual encounters per partner (Nordvik & Liljeros 2006; Chen *et al.* 2008). Therefore, a more important limitation may be our assumptions on how the subpopulations in the MP structure differ from one another. For parsimony, we had assumed these subpopulations to be of the same average size, have the same ratio of men and women, have the same levels of access to care, and have random sexual mixing within subpopulations and similar sexual behaviour other than for the proportion with high sexual risk activity. This, of course, is unrealistic. We believe there are other variations between subpopulations, and some of these may actually facilitate the persistence of gonorrhoea. For instance, NATSAL data revealed that residents of Inner London, an area with higher incidence than the rest of London (Hickman *et al.* 1999; Risley *et al.* 2007), were more likely to have a partner from the same town (i.e. have higher within-subpopulation mixing, π) than individuals resident in the rest of London. Moreover, we also recognize that the interaction between subpopulations is more complex than depicted. For instance, if the subpopulations represented geographically distinct entities, members of certain subpopulations may be more likely to form sexual partnerships with individuals in nearby towns than individuals in towns that are further away (Risley *et al.*

2007). This again is likely to facilitate the persistence of gonorrhoea. It is possible to construct models to depict these details, but such models would require detailed subpopulation level incidence rates and data on sexual behaviour, which are currently not available, and acquisition and analyses of such data could be one area for future research. Were such data available, then it might be possible to construct more detailed MP models that depict key groups of interest, with the choice of such groups driven by the context of the study and the hypotheses under consideration. For instance, empirical data suggest that there may be gender-based differences in the number of partnerships formed by men and women from the same subpopulation, as well in their sexual behaviour and choice of partners. One study of pairs of sexual partners from Baltimore City in the USA showed that women from an at-risk geographical zone (termed 'core area' in that study) were slightly more likely than men to have partners outside the area (Zenilman *et al.* 1999), while the study of case-contact pairs from Colorado Springs (also in the USA) revealed that, while male cases of gonorrhoea mostly reported female partners from the same geographical zone and ethnic group, the female cases tended to report partners of black ethnicity from an at-risk geographical zone (also termed core area in that study). Incorporating gender-based differences in partner selection in a MP model is technically possible if sufficient data are available, and such a study could potentially answer questions about differences in the ratio of male and female cases, help elucidate the role of either gender in transmitting infections between subpopulations, and explain gender-based differences in risk factors for infection. Additional extensions to the MP concept, such as the modelling of dynamics within small geographical zones (e.g. census ward level of high- and low-burden areas in London), and introduction of functional groupings such as sex workers or mobile populations, could be considered if appropriate study questions and data are to surface, and may be possible tracks of investigation for future work. Nevertheless, one final limitation that must be acknowledged is that, while our MP model is better able to replicate observed patterns of gonococcal infection than the AC model against which it is compared, there might be other alternative models and hypotheses which could afford equally or even more plausible explanations for the sociogeographic distribution of gonorrhoea.

For now, however, we believe we have proposed a model that is capable of explaining the unequal distribution of gonococcal cases which has always been apparent in surveillance data (Health Protection Agency 2005), as well as in more sophisticated geographical and molecular epidemiologic analyses (Monteiro *et al.* 2005; Choudhury *et al.* 2006; Risley *et al.* 2007). We also suggest that our MP framework is a more accurate depiction of epidemiological findings supporting the core group hypothesis (Rothenberg 1983; Potterat *et al.* 1985; Zenilman *et al.* 1999; Bernstein *et al.* 2004), and resolves some of the prior ambiguity about what constitutes the core group for STI transmission (Thomas & Tucker 1996). In fact, we would suggest that the entities driving transmission are best referred to as core subpopulations, which we would define as subpopulations that are able to maintain endemic transmission for a particular STI. STIs such as gonorrhoea, which have lower reproductive potential in the presence of adequate access to care, would be restricted to relatively few subpopulations and appear to be unevenly distributed at the population level. On the other hand, STIs such as chlamydia, which have higher reproductive potential, would be spread across more subpopulations and hence appear to be more ubiquitous. Confirming the existence of a MP structure, and identifying the boundaries of the component subpopulations, may be complex, and would probably require a combination of tools—population-based and targeted surveys of sexual behaviour, analysis of detailed surveillance data incorporating sociodemographic and geographical information, and molecular epidemiologic studies. In such a framework, geosocial analyses of incidence rates (e.g. Jennings *et al.* 2005) or molecular epidemiologic data (e.g. Risley *et al.* 2007) could be used to detect and define clusters of transmission and putative core subpopulations. These could then be followed up by field studies to determine and compare the sociodemographic composition and behaviour of these subpopulations with those of non-core subpopulations, to ascertain whether there are indeed higher levels of sexual risk behaviour, or poorer access to STI treatment services in the putative core subpopulations as compared with other population subgroups. For instance, in London, youths of black Caribbean ethnicity living in some inner city areas could be suspected on the basis of incidence, prevalence and molecular epidemiologic analyses to be one of the core subpopulations for gonococcal transmission (Low 2002; Risley *et al.* 2007; Rao *et al.* 2008). Measuring and comparing the health seeking and sexual risk behaviour of this group against that of other subpopulations would allow us to test the hypothesis arising from our MP model that the higher burden of disease in this putative core subpopulation is due to a higher concentration of one or more factors acting at the subpopulation level, such as more risky partnership dynamics, lower use of barrier contraceptives, or delayed access to appropriate treatment. Once identified, such a core subpopulation can be targeted with appropriate and specific interventions such as community-based interventions to change behaviour or screen for STIs (Low 2002; Barlow 2008). These, we propose, are the potential public health implications of the MP framework which make it a subject worthy of further investigation.

REFERENCES

Barlow, D. 2008 Comment: prevalence of *Neisseria gonorrhoeae* infection in young subjects attending community clinics in South London. *Sex. Transm. Infect.* **84**, 121. (doi:10.1136/sti.2007.028423)

Becker, K. M., Glass, G. E., Brathwaite, W. & Zenilman, J. M. 1998 Geographic epidemiology of gonorrhea in Baltimore, Maryland, using a geographic information system. *Am. J. Epidemiol.* **147**, 709–716.

Bernstein, K. T., Curriero, F. C., Jennings, J. M., Olthoff, G., Erbelding, E. J. & Zenilman, J. 2004 Defining core gonorrhea transmission utilizing spatial data. *Am. J. Epidemiol.* **160**, 51–58. (doi:10.1093/aje/kwh178)

Brunham, R. C. & Plummer, F. A. 1990 A general model of sexually transmitted disease epidemiology and its implications for control. *Med. Clin. North Am.* **74**, 1339–1352.

Chen, M. I., Ghani, A. C. & Edmunds, J. 2008 Mind the gap: the role of time between sex with two consecutive partners on the transmission dynamics of gonorrhea. *Sex. Transm. Dis.* **35**, 435–444. (doi:10.1097/OLQ.0b013e3181612d33)

Choudhury, B., Risley, C. L., Ghani, A. C., Bishop, C. J., Ward, H., Fenton, K. A., Ison, C. A. & Spratt, B. G. 2006 Identification of individuals with gonorrhoea within sexual networks: a population-based study. *Lancet* **368**, 139–146. (doi:10.1016/S0140-6736(06)69003-X)

Djuretic, T., Catchpole, M., Bingham, J. S., Robinson, A., Hughes, G. & Kinghorn, G. 2001 Genitourinary medicine services in the United Kingdom are failing to meet current demand. *Int. J. STD AIDS* **12**, 571–572. (doi:10.1258/0956462011923732)

Elliott, L. J., Blanchard, J. F., Beaudoin, C. M., Green, C. G., Nowicki, D. L., Matusko, P. & Moses, S. 2002 Geographical variations in the epidemiology of bacterial sexually transmitted infections in Manitoba, Canada. *Sex. Transm. Infect.* **78**(Suppl. 1), i139–i144.

Farley, T. A., Cohen, D. A. & Elkins, W. 2003 Asymptomatic sexually transmitted diseases: the case for screening. *Prev. Med.* **36**, 502–509. (doi:10.1016/S0091-7435(02)00058-0)

Fenton, K. A. *et al.* 2005 Ethnic variations in sexual behaviour in Great Britain and risk of sexually transmitted infections: a probability survey. *Lancet* **365**, 1246–1255. (doi:10.1016/S0140-6736(05)74813-3)

Foley, E., Patel, R., Green, N. & Rowen, D. 2001 Access to genitourinary medicine clinics in the United Kingdom. *Sex. Transm. Infect.* **77**, 12–14. (doi:10.1136/sti.77.1.12)

Garnett, G. P. 2002 An introduction to mathematical models in sexually transmitted disease epidemiology. *Sex. Transm. Infect.* **78**, 7–12. (doi:10.1136/sti.78.1.7)

Garnett, G. P. & Bowden, J. F. 2000 Epidemiology and control and curable sexually transmitted diseases: opportunities and problems. *Sex. Transm. Dis.* **27**, 588–599. (doi:10.1097/00007435-200011000-00007)

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. *Phil. Trans. R. Soc. B* **354**, 787–797. (doi:10.1098/rstb.1999.0431)

Ghani, A. C. & Garnett, P. G. 2000 Risks of acquiring and transmitting sexually transmitted diseases in sexual partner networks. *Sex. Transm. Dis.* **27**, 579–587. (doi:10.1097/00007435-200011000-00006)

Ghani, A. C., Swinton, J. & Garnett, G. P. 1997 The role of sexual partnership networks in the epidemiology of gonorrhea. *Sex. Transm. Dis.* **24**, 45–56.

Grassly, N. C., Fraser, C. & Garnett, G. P. 2005 Host immunity and synchronized epidemics of syphilis across the United States. *Nature* **433**, 417–421. (doi:10.1038/nature03072)

Griffiths, V., Ahmed-Jushuf, I., Cassell, J. A. & Association for Genitourinary Medicine 2004 Understanding access to genitourinary medicine services. *Int. J. STD AIDS* **15**, 587–589.

Hanski, I. 1998 Metapopulation dynamics. *Nature* **396**, 41–49. (doi:10.1038/23876)

Hardick, J., Hsieh, Y. H., Tulloch, S., Kus, J., Tawes, J. & Gaydos, C. A. 2003 Surveillance of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* infections in women in detention in Baltimore, Maryland. *Sex. Transm. Dis.* **30**, 64–70. (doi:10.1097/00007435-200301000-00013)

Health Protection Agency 2005 *Diagnosis and rates of selected STIs seen at GUM clinics, United Kingdom: 2001–2005.* London, UK: Health Protection Agency.

Heesterbeek, J. A. 2002 A brief history of R_0 and a recipe for its calculation. *Acta Biotheor.* **50**, 189–204. (doi:10.1023/A:1016599411804)

Hethcote, H. W. & Yorke, J. A. 1984 *Gonorrhea transmission, dynamics and control.* Berlin, Germany: Springer.

Hickman, M., Judd, A., Maguire, H., Hay, P., Charlett, A., Catchpole, M., Nayagam, A. & Renton, A. 1999 Incidence of gonorrhoea diagnosed in GUM clinics in South Thames (west) region. *Sex. Transm. Infect.* **75**, 306–311.

Jennings, J., Glass, B., Parham, P., Adler, N. & Ellen, J. M. 2004 Sex partner concurrency, geographic context, and adolescent sexually transmitted infections. *Sex. Transm. Dis.* **31**, 734–739. (doi:10.1097/01.olq.0000145850.12858.87)

Jennings, J. M., Curriero, C. F., Celentano, D. & Ellen, J. M. 2005 Geographic identification of high gonorrhea transmission areas in Baltimore, Maryland. *Am. J. Epidemiol.* **161**, 73–80. (doi:10.1093/aje/kwi012)

Johnson, A. M. *et al.* 2001 Sexual behaviour in Britain: partnerships, practices, and HIV risk behaviours. *Lancet* **358**, 1835–1842. (doi:10.1016/S0140-6736(01)06883-0)

Jungmann, E., Ison, C., Martin, I. M. C. & Fenton, K. A. 2004 Gonorrhoea in young heterosexuals attending urban STD clinics in Britain: a cross-sectional survey. *Int. J. STD AIDS* **15**, 243–248. (doi:10.1258/095646204773557776)

Kerani, R. P., Handcock, M. S., Handsfield, H. H. & Holmes, K. K. 2005 Comparative geographic concentrations of 4 sexually transmitted infections. *Am. J. Public Health* **95**, 324–330. (doi:10.2105/AJPH.2003.029413)

Kolader, M. E., Dukers, H. N., van der Bij, A. K., Dierdorp, M., Fennema, J. S. A., Coutinho, R. A. & Bruisten, S. M. 2006 Molecular epidemiology of *Neisseria gonorrhoeae* in Amsterdam, The Netherlands, shows distinct heterosexual and homosexual networks. *J. Clin. Microbiol.* **44**, 2689–2697. (doi:10.1128/JCM.02311-05)

Kretzschmar, M. 2000 Sexual network structure and sexually transmitted disease prevention: a modeling perspective. *Sex. Transm. Dis.* **27**, 627–635. (doi:10.1097/00007435-200011000-00011)

Lacey, C. J., Merrick, D. W., Bensley, D. C. & Fairley, I. 1997 Analysis of the sociodemography of gonorrhoea in Leeds, 1989–93. *Br. Med. J.* **314**, 1715–1718.

Lavelle, S. J., Jones, K. E., Mallinson, H. & Webb, A. M. C. 2006 Finding, confirming, and managing gonorrhoea in a population screened for chlamydia using the Gen-Probe Aptima Combo2 assay. *Sex. Transm. Infect.* **82**, 221–224. (doi:10.1136/sti.2005.017616)

Lin, J. S. *et al.* 1998 Transmission of *Chlamydia trachomatis* and *Neisseria gonorrhoeae* among men with urethritis and their female sex partners. *J. Infect. Dis.* **178**, 1707–1712. (doi:10.1086/314485)

Low, N. 2002 Phase specific strategies for the prevention, control, and elimination of sexually transmitted infections: case study in Lambeth, Southwark, and Lewisham, London, UK. *Sex. Transm. Infect.* **78**(Suppl. 1), i133–i138.

Low, N., Daker-White, G., Barlow, D. & Pozniak, A. L. 1997 Gonorrhoea in inner London: results of a cross sectional study. *Br. Med. J.* **314**, 1719–1723.

Monteiro, E. F., Lacey, C. J. & Merrick, D. 2005 The interrelation of demographic and geospatial risk factors between four common sexually transmitted diseases. *Sex. Transm. Infect.* **81**, 41–46. (doi:10.1136/sti.2004.009431)

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. (doi:10.1097/01.olq.0000194601.25488.b8)

Potterat, J. J., Rothenberg, R. B., Woodhouse, D. E., Muth, J. B., Pratts, C. I. & Fogle II, J. S. 1985 Gonorrhea as a social disease. *Sex. Transm. Dis.* **12**, 25–32.

Rao, G. G., Bacon, L., Evans, J., Dejahang, Y., Michalczyk, P. & Donaldson, N. 2008 Lewisham chlamydia and gonoccoccus screening programme. Prevalence of *Neisseria gonorrhoeae* infection in young subjects attending community clinics in South London. *Sex. Transm. Infect.* **84**, 117–121. (doi:10.1136/sti.2007.026914)

Risley, C. L., Ward, H., Choudhury, B., Bishop, C. J., Fenton, K. A., Spratt, B. G., Ison, C. A. & Ghani, A. C. 2007 Geographical and demographic clustering of gonorrhoea in London. *Sex. Transm. Infect.* **83**, 481–487. (doi:10.1136/sti.2007.026021)

Rothenberg, R. B. 1983 The geography of gonorrhea. Empirical demonstration of core group transmission. *Am. J. Epidemiol.* **117**, 688–694.

Rothenberg, R. B. & Potterat, J. J. 1988 Temporal and social aspects of gonorrhea transmission: the force of infectivity. *Sex. Transm. Dis.* **15**, 88–92.

Schneeberger, A., Mercer, C. H., Gregson, S. A. J., Ferguson, N. M., Nyamukapa, C. A., Anderson, R. M., Johnson, A. M. & Garnett, G. P. 2004 Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. *Sex. Transm. Dis.* **31**, 380–387. (doi:10.1097/00007435-2004 06000-00012)

Shahmanesh, M., Gayed, S., Ashcroft, M., Smith, R., Roopnarainsingh, R., Dunn, J. & Ross, J. 2000 Geomapping of chlamydia and gonorrhoea in Birmingham. *Sex. Transm. Infect.* **76**, 268–272. (doi:10.1136/sti.76.4.268)

Thomas, J. C. & Tucker, M. J. 1996 The development and use of the concept of a sexually transmitted disease core. *J. Infect. Dis.* **174**(Suppl. 2), S134–S143.

Thomas, R. & Smith, T. E. 2000 Multiregion contact systems for modelling STD epidemics. *Stat. Med.* **19**, 2479–2491. (doi:10.1002/1097-0258(20000915/30)19:17/18<2479:: AID-SIM583>3.0.CO;2-0)

Turner, K. M., Garnett, G. P., Ghani, A. C., Sterne, J. A. C. & Low, N. 2004 Investigating ethnic inequalities in the incidence of sexually transmitted infections: mathematical modelling study. *Sex. Transm. Infect.* **80**, 379–385. (doi:10. 1136/sti.2003.007575)

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. (doi:10.1186/ 1742-4682-3-3)

Wasserheit, J. N. & Aral, S. O. 1996 The dynamic topology of sexually transmitted disease epidemics: implications for prevention strategies. *J. Infect. Dis.* **174**(Suppl. 2), S201–S213.

White, P. J., Ward, H., Cassell, J. A., Mercer, C. H. & Garnett, G. P. 2005 Vicious and virtuous circles in the dynamics of infectious disease and the provision of health care: gonorrhea in Britain as an example. *J. Infect. Dis.* **192**, 824–836. (doi:10.1086/432004)

Zenilman, J. M., Ellish, N., Fresia, A. & Lass, G. 1999 The geography of sexual partnerships in Baltimore: applications of core theory dynamics using a geographic information system. *Sex. Transm. Dis.* **26**, 75–81. (doi:10.1097/00007435-199902000-00002)