

Estimating the sexual mixing patterns in the general population from those in people acquiring gonorrhoea infection: theoretical foundation and empirical findings

Adrian Renton, Luke Whitaker, Cathy Ison, Jane Wadsworth, J R W Harris

Abstract

Study objectives - To describe mathematically the relationship between patterns of sexual mixing in the general population and those of people with gonorrhoea infection, and hence to estimate the sexual mixing matrix for the general population.

Design - Integration of data describing sexual behaviour in the general population, with data describing sexual behaviour and mixing among individuals infected with gonorrhoea. Use of these data in a simple mathematical model of the transmission dynamics of gonorrhoea infection.

Setting - The general population of London and a genitourinary medicine (GUM) clinic in west London.

Participants - These comprised 1520 men and women living in London who were randomly selected for the national survey of sexual attitudes and lifestyles and 2414 heterosexual men and women who presented to the GUM clinic with gonorrhoea.

Main results - The relationship between sexual mixing among people with gonorrhoea and sexual mixing in the general population is derived mathematically. An empirical estimate of the sexual mixing matrix for the general population is presented. The results provide tentative evidence that individuals with high rates of acquisition of sexual partners preferentially select other individuals with high rates as partners (assortative mixing).

Conclusions - Reliable estimates of sexual mixing have been shown to be important for understanding the evolution of the epidemics of HIV infection and other sexually transmitted diseases. The possibility of estimating patterns of sexual mixing in the general population from information routinely collected in gonorrhoea contact tracing programmes is demonstrated. Furthermore, the approach we describe could, in principle, be used to estimate the same patterns of mixing, using contact tracing data for other sexually transmitted diseases, thus providing a way of validating our results.

Mathematical models of the transmission dynamics of sexually transmitted disease (STD) are increasingly being used to predict trends in the occurrence of these diseases, and to inform control strategies. These models require information about rates of partner change and patterns of sexual mixing in populations. The national survey of sexual attitudes and lifestyles (NSSAL) provides us with good information about rates of partner change within the UK population but tells us little about patterns of sexual mixing, which are difficult to estimate directly. In this paper we present data which we have obtained about recent sexual behaviour in patients with gonorrhoea attending St Mary's Hospital, Paddington, including pairs of individuals between whom gonorrhoea transmission is thought to have occurred. Information from the NSSAL for respondents living in London is also presented to provide estimates of the distribution of sexual contact rates in the general population from which individuals with gonorrhoea infection are drawn. Using a deterministic model of the transmission dynamic of gonorrhoea infection we derive an expression which allows us to estimate the sexual mixing matrix for the general population from these data. We present the results and discuss the utility of our approach.

Introduction

The emergence of the pandemic of HIV infection over the past 10 years has stimulated interest in achieving a greater understanding of sexual behaviour and the epidemiology and transmission dynamics of sexually transmitted infections. In this context trends in the occurrence of postpubertal gonococcal infection have received considerable attention as a potential indicator of underlying sexual behaviour¹⁻³ in populations, and the Government's health strategy, *The Health of the Nation* has set a specific target of a 20% reduction, from the 1990 level, in the incidence of this infection by 1995.⁴ Despite this attention, the relationship between gonorrhoea incidence and underlying patterns of sexual behaviour in populations remains poorly understood.

Mathematical models provide a way of formalising assumptions about the relationship between patterns of sexual behaviour and the occurrence of sexually transmitted infections, and allow predictions based on these assumptions. Early work by Hethcote and Yorke⁵

Academic Department
of Public Health,
Medicine
A Renton
L Whitaker
J Wadsworth

Department of
Bacteriology,
C Ison

and Department of
Genitourinary
Medicine,
JRW Harris

St Mary's Hospital,
Norfolk Place,
London W2 1PG

Correspondence to:
Dr A Renton.

Accepted for publication
October 1994

showed that deterministic models of the transmission dynamics of gonorrhoea could be constructed using a limited number of parameters summarising the pertinent features of both the natural history of the host-gonococcus interaction and the patterns of sexual behaviour occurring in populations. These authors demonstrated that the key parameters in natural history which determine the transmission dynamic are the probabilities of transmission between an infected and uninfected individual within a partnership (*transmission probability*) and the period of time for which infected individuals remain infected, or its reciprocal (*the recovery rate*).⁵ To model the influence of patterns of sexual behaviour, they divided the population into strata of varying size, defined both by gender and by specified average rates of acquisition of new partners (*sexual contact rate*) pertaining to each stratum. They further defined the probabilities that new partners acquired by individuals within each stratum would themselves be drawn from any particular stratum. The resulting matrix of probabilities is termed the contact or *mixing matrix*. We have previously shown that these models may have profound implications for the interpretation of secular trends in the occurrence of gonorrhoea.⁶

The simplest assumption which can be made about the mixing matrix is that of proportional or random mixing. Under this assumption the probability that new partners acquired by stratum *i* are members of stratum *j* is equal to the proportion of all potential partnerships available to stratum *i* which are provided by stratum *j*. This proportion is, in turn, determined by the proportion of the total population who are in stratum *j* and their relative rate of acquisition of new partners. Actual mixing patterns may vary, either towards assortative mixing (where those within a particular sexual contact rate stratum preferentially select as partners others from within that stratum, beyond the extent expected through proportional mixing) to disassortative mixing (where those within a particular sexual contact rate stratum preferentially select as partners others from within different strata). Using a proportional mixing matrix in their model, Hethcote and Yorke were able to write differential equations describing the rates of change of numbers of infected and uninfected individuals within theoretical populations, and to solve these equations for the endemic equilibrium case in which these numbers are constant.⁵ Thus the equilibrium prevalence and incidence of gonococcal infections in populations can be calculated where the infection transmission probabilities, recovery rates, and distributions of sexual contact rates are known.

More recently, there has been growing interest in using such models to predict trends in the epidemic of HIV infection in different populations.⁸⁻¹² Anderson *et al* and others have elaborated deterministic models which describe the dynamics of sexual transmission of HIV infection and have shown that deviations from proportional mixing either to the assortative (like with like) or disassortative (like with unlike) extremes have the potential to exert

a major influence on the transmission dynamic.⁸⁻¹¹

There has been particular interest in exploring whether, under different assumptions about the transmission probabilities and patterns of infectiousness of HIV, a heterosexual epidemic could be sustained. Attempts in this area have been hampered by the lack of availability of descriptive information on the sexual contact rates and sexual mixing patterns within heterosexual populations of interest. In the UK, information about sexual contact rates has been provided with the publication of the findings of the NSSAL.¹³ However, empirical information on the population mixing matrix remains elusive. Direct measurement of these patterns through sampling and surveying the general population presents methodological and ethical difficulties. Biases could be introduced by non-disclosure and by under sampling of epidemiologically important but hard to reach populations. Ethical problems arise as direct measurement would require disclosure by one individual of sensitive information about others, without the consent of those others.⁷

In the face of these problems indirect approaches to estimating sexual mixing patterns have been considered. In particular the potential to exploit situations in which contact tracing is, in any case, carried out, such as where individuals present to clinical services with sexually transmitted infections has been proposed by Garnett and Anderson.⁷ These authors have made a preliminary attempt to take this approach using information obtained from the contact tracing of individuals with gonorrhoea in New York, previously published by Rothenburg,¹⁴ and similar data from Colorado Springs, previously published by Potteratt *et al*.¹⁵ Available information on contact pairs was, however, restricted to place of residence, and mixing matrices and a parameter summarising the degree of disassortativeness of mixing were derived for strata based on residence in areas of high, medium, and low gonorrhoea incidence. The authors acknowledge some of the limitations of this geographically based approach and note the desirability of obtaining contact tracing information for individuals with STDs which allows the construction of mixing matrices based on compartmentalisation into strata defined by sexual contact rate. Furthermore, the authors point out that even if the sexual mixing patterns of people with an STD were known, these patterns could not be taken as representative of the mixing patterns of the population at large. This is because individuals with high sexual contact rates, or who preferentially select partners who have high contact rates are likely to be at increased risk of STD. Thus, individuals with STD will be a biased sample of all individuals in the population, specifically with regard to their sexual contact rates and sexual mixing preferences. Garnett and Anderson suggest, therefore, that there is a need to develop methods of inference which will allow a general description of parameters defining sexual mixing in populations from incomplete samples. It is this problem which we address

Table 1 Observed distribution of heterosexual men and women attending the clinic with gonorrhoea in relation to numbers of sexual partners in the past month

Gender	No of opposite sex partners in past month				Total
	1	2	3	4	
Men	820	591	64	10	1485
Women	788	141	0	0	929

in this paper. Granath *et al*¹⁶ have developed a different approach to our own, based on a theoretical derivation of conditional probabilities of infection with *Chlamydia trachomatis* within different sexual contact rate strata. They used contact tracing data for Chlamydia to derive a population mixing matrix which showed significant deviation from proportional mixing. Their approach, however, is problematic because it seems to assume a constantly increasing prevalence, although this is not made clear in their discussion.

In this paper we discuss a novel approach to estimating the population sexual mixing matrix. We present data which we have obtained about recent sexual behaviour in patients with gonorrhoea² attending St Mary's Hospital, Paddington, including pairs of individuals between whom gonorrhoea transmission is thought to have occurred. Information from the NSSAL¹³ for respondents living in London is also presented to provide estimates of the distribution of sexual contact rates in the population from which individuals and pairs with gonorrhoea infection are drawn. We derive an expression by which the patterns of sexual mixing of individuals who attend STD clinics with gonorrhoea infection and their contacts can be used to infer the sexual mixing patterns of the wider population. This is achieved by appropriate manipulation of a deterministic model similar to that presented by Garnett and Anderson,¹² which is based on the original Hethcote Yorke model.⁵ We use estimates of the parameters and variables in the model derived from the empirical data to estimate the population sexual mixing matrix. We assess the sensitivity of the results to errors in the empirical estimates of the parameters and variables, and discuss the utility of our approach.

The data

SEXUAL MIXING OF INDIVIDUALS WITH GONORRHOEA INFECTION

We have previously published data from our study of the occurrence of gonorrhoea and the sexual behaviour of infected individuals at the genitourinary medicine (GUM) clinic at St

Table 2 Observed distribution of heterosexual transmission pairs classified in relation to number of sexual partners in the past month of both male and female partner

	No of male sexual partners during past month of woman in transmission pair	
	1	2
No of female sexual partners during last month of man in transmission pair	1	83
	2	72
	3	12
	4	2
		14
		11
		1
		0

Mary's Hospital, Paddington.^{2,17,18} All patients from whom *Neisseria gonorrhoeae* was isolated by the diagnostic laboratory serving the GUM clinic between January 1988 and December 1993 (the study period) were identified by hospital number. Clinic notes were sought, and found in 97% of a total of 3363 cases. Details of age and HIV serostatus were recorded from the notes for each visit at which the organism was isolated.

When an individual (*an index case*) is diagnosed as having a gonococcal infection at the clinic he or she receives immediate treatment and interview by a specially trained health adviser. The health adviser counsels the individual and obtains a recent sexual history, noting the number and gender of all persons (*contacts*) with whom the index case has had sexual relations during the month before presentation and the type of intercourse which occurred. Contact slips are issued recording the index case's clinic number and a coded diagnosis and advice on where and how to seek medical attention. The *index case* is asked to pass one slip on to each contact. Where a contact subsequently presents to the same clinic, a record of this contact's clinic number and any diagnosis is made in the notes of the index case. Where the contact is found to be infected with gonorrhoea, he or she will be referred to the health adviser for tracing of contacts in turn.

Patient based information on the number of same sex and opposite sex contacts during the month before presentation were obtained from the detailed records kept by health advisors. Individuals were defined as heterosexual where they had had only opposite sex partners during the previous month. Altogether 1485 heterosexual men (median age = 26 years, range 15-72) and 929 heterosexual women (median age = 22 years, range 14-53) were identified. The distribution of the whole sample of heterosexual men and women attending with gonorrhoea infection during the study period across the sexual contact rate strata defined by the number of partners in the last month is shown in table 1.

The health adviser records also allow the identification of series of contact pairs. We have previously shown very high concordance between the serovar/auxotype group of strains isolated from such contact pairs¹⁸ which provides strong evidence that transmission between individuals in these pairs has in fact occurred. Therefore, we infer that for health adviser contact pairs where both individuals have had gonorrhoea infection, transmission has occurred and term these *gonorrhoea transmission pairs*. All known heterosexual transmission pairs, totalling 195 attending the clinic during the study period were identified. These 195 transmission pairs are distributed into classes defined by sexual contact rate strata of both partners in table 2.

SEXUAL MIXING IN THE LONDON HETEROSEXUAL POPULATION

Information about the number of heterosexuals' sexual partners in the last month is

Table 3 Observed distribution of men and women in the London sample of NSSAL in relation to numbers of sexual partners in past month

Gender	No of opposite sex partners in past month		
	1	2	3
Men	685	18	3
Women	796	18	0

also available from the NSSAL. This study included a random sample of men and women aged 16–59 years living in Great Britain. The distribution of men and women living in Greater London who had had only same sex partners in the last month from this sample is shown in table 3.

The model

Having identified and presented the empirical data which are available to us, we now develop expressions for the elements of the mixing matrix for the wider population from which the individuals presenting to our clinic with gonorrhoea infection are drawn. We follow the deterministic compartmental framework for a heterosexual model outlined by Garnett and Anderson¹² following Hethcote and York. We have, however, dropped the assumption of movement of individuals in and out of the sexually active population for simplicity's sake. We have revised our notation from our previous papers to be consistent with that used by Garnett and Anderson.

THE BASIC MODEL

The population comprises numbers of both susceptible uninfected individuals (*uninfecteds*) (X_{ki}) and infectious, infected non-susceptible individuals (*infecteds*) (Y_{ki}) of sex k and sexual contact rate stratum i , where contact rate stratum is defined by the rate of acquisition of new sexual partners (sexual contact rate). The total sexually active population is taken to be of size N , and the total number of individuals of sex k in activity class i is denoted N_{ki} . The parameter β_k represents the probability of transmission from sex k' to sex k per new contact. The parameter c_{ki} is the rate of acquisition of new sexual partners (sexual contact rate) of sex k in activity stratum i . The parameter ρ_{ki} represents the proportion of new sexual partners acquired by sex k in activity stratum i , which are of the opposite sex k' and activity class j . The parameters ρ_{kij} are the elements of the mixing matrix for the population. The parameter v_k is defined as the rate of recovery from infection in sex k . Two constraints are placed on the system of parameters:

$$\sum_{j=1}^n \rho_{kij} = 1 \quad (1)$$

where n is the number of activity classes, and further that:

$$\rho_{kij} c_{ki} N_{ki} = \rho_{k'ji} c_{k'j} N_{k'j} \quad (2)$$

The model gives a system of differential equations for the rate of change of the number of infecteds (\dot{Y}_{ki}) and uninfecteds (\dot{X}_{ki}) as:

$$\dot{Y}_{ki} = \beta_k c_{ki} X_{ki} \sum_{j=1}^n \rho_{kij} \frac{Y_{k'j}}{N_{k'j}} - v_k Y_{ki} \quad (3)$$

and

$$\dot{X}_{ki} = v_k Y_{ki} - \beta_k c_{ki} X_{ki} \sum_{j=1}^n \rho_{kij} \frac{Y_{k'j}}{N_{k'j}} \quad (4)$$

Where the infection is in endemic equilibrium, which we define to mean that $\dot{Y}_{ki} = \dot{X}_{ki} = 0$ for all activity class/gender combinations, then the incidence is constant over time and equal to the recovery rate. Thus we can write from (4):

$$Y_{ki} = \frac{1}{v_k} \beta_k c_{ki} X_{ki} \sum_{j=1}^n \rho_{kij} \frac{Y_{k'j}}{N_{k'j}} \quad (5)$$

where Y_{ki} is the number of infecteds at equilibrium. (Henceforth all terms Y , X , and N will be used to denote the values of the variables at equilibrium.)

EXTENDING THE BASIC MODEL

As we have explained in the introduction, we want to estimate the mixing matrix ρ for the population. We presented data in table 2 which describes mixing patterns in gonorrhoea transmission pairs, but mixing among these people is likely to be quite different from that in the wider population. We will extend the basic model in order to allow estimation of ρ from these data.

From equation (5) we can express the number of prevalent cases occurring in stratum ki which were acquired through contact with stratum $k'j$. We denote this number ${}_{k'j}Y_{ki}$. Thus

$${}_{k'j}Y_{ki} = \frac{1}{v_k} \beta_k c_{ki} X_{ki} \rho_{kij} \frac{Y_{k'j}}{N_{k'j}} \quad (6)$$

We define a prevalent gonorrhoea transmission pair as an individual with current infection, paired with the sexual partner from whom the infection was acquired. If all transmission pairs in the population were known at any time, then the number of these pairs which comprised an individual in stratum ki together with an individual in stratum $k'j$ (which we will denote $M_{k'jik}$) is simply equal to the sum of the number of cases occurring in stratum ki which were acquired through contact with stratum $k'j$, and the number of cases occurring in stratum $k'j$ which were acquired through contact with stratum ki . These are given in (6) as ${}_{k'j}Y_{ki}$ and ${}_{ki}Y_{k'j}$ respectively. Thus:

$$M_{k'jik} = {}_{k'j}Y_{ki} + {}_{ki}Y_{k'j} \quad (7)$$

$$M_{k'jik} = \frac{1}{v_k} \beta_k c_{ki} X_{ki} \rho_{kij} \frac{Y_{k'j}}{N_{k'j}} + \frac{1}{v_{k'}} \beta_{k'} c_{k'j} X_{k'j} \rho_{k'ji} \frac{Y_{ki}}{N_{ki}} \quad (8)$$

now from (2) we have $\rho_{k'ji} = \frac{\rho_{kij} c_{ki} N_{ki}}{c_{k'j} N_{k'j}}$ therefore

$$M_{k'jik} = \frac{1}{v_k} \beta_k c_{ki} X_{ki} \rho_{kij} \frac{Y_{k'j}}{N_{k'j}} + \frac{1}{v_{k'}} \beta_{k'} c_{k'j} X_{k'j} \rho_{kij} \frac{Y_{ki}}{N_{k'j}} \quad (9)$$

which we can rewrite as:

$$M_{k'ik'j} = \rho_{k'ij} \frac{c_{ki}}{N_{k'j}} X_{ki} Y_{ki} \left(\frac{\beta_k}{v_k} \cdot \frac{Y_{k'j}}{Y_{ki}} + \frac{\beta_{k'}}{v_{k'}} \cdot \frac{X_{k'j}}{X_{ki}} \right) \tag{10}$$

Estimating the mixing matrix for the wider population from the empirical data

ESTIMATING THE PARAMETERS AND VARIABLES
 Examination of equation (10) allows us to identify the parameters and variables which must be estimated empirically to permit the solution of this system of equations for the parameters $\rho_{k'ij}$ for sex $k=m$ as men and $k'=f$ as women. All rates are defined as per month. Average estimates of β_m and β_f , v_m and v_f are taken from the literature²⁰ as $\beta_m = 0.6$, $\beta_f = 0.8$, $v_m = v_f = 1$ (that is mean duration of infectivity in men and women is one month, which we believe is a reasonable estimate in our population). We need estimates of the values M_{mif} . In addition we need information about terms involving Y , X , and N for people of both sexes and all contact rate strata. The sensitivity of the terms $\rho_{k'ij}$ to variation in the values of these parameters is considered in the appendix. We have taken c_i as the number of sexual contacts in the last month as reported to the health advisor or the NSSAL.

Estimating mixing in transmission pairs (M_{mif})

Assuming that the gonorrhoea transmission pairs, identified through clinic contact tracing in our GUM clinic, are a random sample of all such transmission pairs within the population (M) from which cases are drawn, then our observed numbers of pairs (which we denote m_{mif}) which are shown in table 2 will be a fraction of M_{mif} which is constant across all m and f . Thus:

$$\alpha m_{mif} = M_{mif} \tag{11}$$

where α is reciprocal of the sampling fraction. Thus from (10) we have:

$$\alpha m_{mif} = \rho_{mif} \frac{c_{mi}}{N_{fi}} X_{mi} Y_{mi} \left(\frac{\beta_m}{v_m} \cdot \frac{Y_{fi}}{Y_{mi}} + \frac{\beta_f}{v_f} \cdot \frac{X_{fi}}{X_{mi}} \right) \tag{12}$$

Estimating terms involving Y, X, and N

The data relating to cases of gonorrhoea which we presented in section 2 are based upon attendances at a single GUM clinic. Because we do not have adequate information on the size of the population in which these cases are occurring, we cannot estimate X and N . However, it will be seen in the next section that we can solve (12) for ρ using only estimates of the ratios (Y_{mi}/Y_{fi}) , (X_{mi}/X_{fi}) and (N_{mi}/N_{fi}) .

The distribution of the whole sample of men and women attending with gonorrhoea in-

fection between January 1988 and December 1993 across sexual contact rate strata is shown in table 1 above. If we denote the elements of this table y_{mi} , y_{fi} and assuming that these cases are an unbiased sample of all such cases prevalent in the population at any time, then (Y_{mi}/Y_{fi}) is estimated by (y_{mi}/y_{fi}) . Stratum i therefore consists of individuals who have had i sexual partners in the past month.

The relative size of the sexual contact rate strata in the population from which the gonorrhoea cases are drawn (N_{mi}/N_{fi}) can be directly estimated from the London sample of the NSSAL. The distribution of the entire London sample is shown in table 3 above. Since the prevalence of gonorrhoea may be expected to be small, even in the high contact rate strata, (N_{mi}/N_{fi}) can be taken as a reasonable estimate of (X_{mi}/X_{fi}) in all cases. We have applied a constant multiplier of 1.14 to these ratios to make the total number of sexual contacts consistent between men and women.

ESTIMATING THE MIXING MATRIX

Given the estimates of the parameters and variables which we have made, the system of equations (12) reduces to a system of $i \times j$ equations in $i \times j + 1$ unknowns, that is, ρ_{mif} and α . However, equation (1) allows us to reduce the number of unknowns by i , leaving the equations overspecified.

We can easily solve the equation (12) for ρ_{mif} for the case where we have only two sexual contact strata among women. If we write equation (12) separately for sexual contact pairs comprising men in stratum i , and women in strata 1 and 2 we have:

$$\alpha m_{mif1} = \rho_{mif1} \frac{c_{mi}}{N_{f1}} X_{mi} Y_{mi} \left(\frac{\beta_m}{v_m} \cdot \frac{Y_{f1}}{Y_{mi}} + \frac{\beta_f}{v_f} \cdot \frac{X_{f1}}{X_{mi}} \right) \tag{13}$$

$$\alpha m_{mif2} = \rho_{mif2} \frac{c_{mi}}{N_{f2}} X_{mi} Y_{mi} \left(\frac{\beta_m}{v_m} \cdot \frac{Y_{f2}}{Y_{mi}} + \frac{\beta_f}{v_f} \cdot \frac{X_{f2}}{X_{mi}} \right) \tag{14}$$

Dividing (13) by (14) and eliminating α , X_{mi} and Y_{mi} and substituting $\rho_{mif2} = (1 - \rho_{mif1})$ (because there are only two female sexual contact rate strata) gives:

$$\rho_{mif1} = \frac{1}{\frac{m_{mif2}}{m_{mif1}} \cdot \frac{N_{f2}}{N_{f1}} \left(\frac{\beta_m}{v_m} \cdot \frac{Y_{f1}}{Y_{mi}} + \frac{\beta_f}{v_f} \cdot \frac{X_{f1}}{X_{mi}} \right) \left(\frac{\beta_m}{v_m} \cdot \frac{Y_{f2}}{Y_{mi}} + \frac{\beta_f}{v_f} \cdot \frac{X_{f2}}{X_{mi}} \right) + 1} \tag{15}$$

Then all the remaining elements ($\rho_{k'ij}$) of the population mixing matrix can be calculated from equations (1) and (2).

Table 4 Estimates of the elements of the mixing matrix and expected values on proportional mixing assumption for male and female sexual activity classes

ρ	Empirical estimate	Expected value on proportional mixing assumption
Men:		
ρ_{m11}	0.96	0.96
ρ_{m12}	0.04	0.04
ρ_{m21}	0.88	0.96
ρ_{m22}	0.12	0.04
ρ_{m31}	0.94	0.96
ρ_{m32}	0.06	0.04
Women:		
ρ_{f11}	0.94	0.94
ρ_{f12}	0.05	0.05
ρ_{f13}	0.01	0.01
ρ_{f21}	0.91	0.94
ρ_{f22}	0.13	0.05
ρ_{f23}	0.02	0.01

Stratum i consists of individuals who have had i sexual partners in the last month.

We have used the empirical data presented above to solve equation (15) and have thence derived an estimate of the elements ρ_{mij} which form the mixing matrix of the population from which the gonorrhoea cases are drawn. Stratum i consists of individuals who have had i sexual partners in the past month. This matrix is shown in table 4.

It is important to note that because these estimates are derived from overspecified equations and empirical data the results are not guaranteed to be self consistent (that is, an exact solution may not be possible). There are a number of ways that we could have chosen to solve the equations, and these will result in a family of different solutions. The extent to which these solutions differ will reflect the extent of inconsistency in the equations.

Our method for the solution of these overspecified equations was chosen for algebraic convenience. We took the pairs of equations (13) and (14) for $i=1, 2$, and 3 , and, having substituted ρ_{m11} for ρ_{m12} in (14), eliminated α independently for each level of i . A consequence of the overspecification is that the estimate of α we obtain by substituting ρ_{m11} and ρ_{m12} back into (13) or (14) is different for each of these levels. An additional consequence is that $\sum_j \rho_{ij}$ is not guaranteed to be unity as required by equation (1).

To assess the extent of any inconsistency we have calculated values of α (the reciprocal of the notional sampling fraction of transmission pairs) by back substitution, as explained above, and obtained values 12.9, 10.5, and 12.5. The reasonable agreement between these three values, together with the fact that our estimates of ρ_{ij} given in table 5 are reasonably consistent with equation (1), leads us to believe that the family of possible solutions will all be similar.

There is tentative evidence of assortative mixing in the estimates of ρ for the men and women who have had more than one sexual partner in the past month. This is shown by the fact that the ρ_{m21} and ρ_{f21} are smaller than expected from proportional mixing, while ρ_{m22} and ρ_{f22} are larger than expected. A discussion of the sensitivity of ρ both to errors in our estimates of the model parameters and to sampling error in our observed data is provided

in the appendix, together with estimates of confidence intervals of ρ .

Discussion

We believe that our findings represent the first report of an attempt to derive a sexual mixing probability matrix for the wider population from which individuals with gonorrhoea infection are drawn, directly from information derived from contact tracing programmes. We have elaborated a theoretical framework for accomplishing this and shown that the results are both plausible and consistent when the framework is applied to real data derived from both routinely collected contact tracing data and a population based sexual behaviour survey. We believe that the approach we have taken is applicable to endemic sexually transmitted diseases in general. However, the adoption of our combined theoretical/empirical approach demands close scrutiny of both theoretical assumptions and potential biases in the empirical estimates of the variables and parameters of the model.

The key assumptions of our use of the Yorke Hethcote model are that gonorrhoea infection is in a state of endemic equilibrium over time. That is to say that the overall incidence rate of infection is equal to the rate of recovery of infected cases. Furthermore, it is assumed that the incidence rates within the individual sexual contact rate strata are also stable over time and that the system is non-oscillatory. An additional assumption is that the transmission probabilities and recovery rates for men and women are constant over sexual contact rate strata and over time. Over the period for which we have collected data, the rates of presentation of gonorrhoea cases to our own clinic and to others in the area have remained relatively stable, as has the distribution of cases among different contact strata. While we have no information concerning the stability of the transmission probabilities over the period of observation, there have been some changes in the distribution of gonococcal isolates across serological types of gonococcus which emphasise the need for caution. Furthermore, changes in condom usage over time might have had a significant influence on transmission probability. However, our estimates of population mixing matrices are not particularly sensitive to changes in this parameter. The model further assumes that all individuals within a particular sexual contact rate class have the same probability of making contact with an individual from any particular sexual contact rate class of the opposite sex. This assumption may well be incorrect. However, further attempts to measure sexual networks directly will be required before this can be adequately assessed.

To derive our estimates of the mixing probabilities we have used empirical data drawn from contact tracing information collected by health advisers. The potential biases in these data have been well summarised by Granath *et al*¹⁶ and we will focus here on the particular problems which relate to our own data. One key assumption which we have made is that

the *individual cases* of gonorrhoea which present to our clinic are a random sample of all such cases within the population and that they are distributed across sexual activity classes in exactly the same proportions as for all cases in the population served. In addition, we have assumed that the gonorrhoea *transmission pairs* which we have identified represent a random sample of all such transmission pairs from which the cases are drawn. While we feel reasonably comfortable with the first of these assumptions, there are *a priori* reasons to suppose that the second may not be correct. In particular, we may have been more likely to have identified contact pairs where both individuals would naturally tend to attend our clinic due to the proximity of their homes or work, or where individuals fall into lower contact rate strata, because they are less likely to have attended elsewhere already as a result of contact tracing from further up the chain of transmission. We have no direct evidence that this is the case, but are seeking to compare characteristics of our attenders and transmission pairs with those of other clinics in the area to investigate this further.

We have used data drawn from the London sample of the NSSAL to estimate the distribution, across sexual contact rate classes, of individuals in the population from which cases are drawn. The fieldwork for the survey took place exactly around the midpoint of our period of observation. However, the sample was selected so as to be representative of the London population as a whole. While the catchment population of our clinic is difficult to define, the population of the area in which the clinic is located differs demographically from the NSSAL sample. Our findings are quite sensitive to differences in the sexual contact rate distribution in the population and this is a shortcoming of the results which we present.

A problem arises in the application of the model when parameters from real people are applied to it. We have treated the sexual contact rate as identical to the number of sexual partners which an individual has in a defined time period, in this case one month. There is, however, an implicit assumption in the model that the sexual contact rate should include only the rate of acquisition of new partners who have the possibility of being infected. Our operationalisation of the sexual contact rate is therefore problematic in that individuals who have had, for instance, only one sexual partner may have acquired no new partners during the period. Then their risk of infection is determined by the number of sexual contacts of their partner in a defined period. This raises the question as to how a parameter for the effective contact rate might be operationalised. We are currently attempting to achieve this using a model which distinguishes regular and casual or new contacts.

Despite these problems in sampling and definition, we believe that our approach has the potential to be used more generally in attempting to study sexual mixing patterns and the transmission dynamics of STDs. Our preliminary results suggest some degree of as-

sortativeness in the population sexual mixing patterns. The strength of our approach is that it can be used for a range of different STDs to arrive at the same mixing matrix, thus providing the potential for the validation of many of the assumptions which we have made. Therefore, we are currently collecting contact tracing data for *Chlamydia trachomatis* from the same clinic, and hope to use the same approach to validate our findings from the gonorrhoea data in the same population.

The value of mathematical modelling approaches to understanding the transmission dynamics of STDs is crucially dependent on an adequate knowledge of the sexual mixing patterns within the general population. STD contact tracing information provides an approach to the indirect measurement of these patterns. Immediate priorities for research include studies to estimate reliably the transmission probabilities and recovery rates for different STDs, and more detailed exploration of the sexual networks of infected and uninfected individuals. As such work progresses, we should arrive at a point where routine contact tracing information combined with the analysis of trends in the occurrence of STDs will provide us with a detailed picture of the sexual behaviour of populations, and the potential to predict better new epidemics and to develop more effective control strategies.

We would like to thank Alan Giles for his tremendous clerical efforts without which this work would not have been possible.

Appendix

SENSITIVITY ANALYSIS AND CONFIDENCE INTERVALS ON ρ

To assess the sensitivity of our estimates of the elements ρ_{kik_j} of the mixing matrix given in equation (15) we first write an expression for the partial derivatives with respect to an arbitrary parameter x , where x represents any one of the parameters or variables in equation (15) which we repeat here.

$$\rho_{mi1} = \frac{1}{\frac{m_{mif2}}{m_{mif1}} \cdot \frac{N_{f2}}{N_{f1}} \left(\frac{\beta_m}{v_m} \cdot \frac{Y_{f1}}{Y_{mi}} + \frac{\beta_f}{v_f} \cdot \frac{X_{f1}}{X_{mi}} \right) \left(\frac{\beta_m}{v_m} \cdot \frac{Y_{f2}}{Y_{mi}} + \frac{\beta_f}{v_f} \cdot \frac{X_{f2}}{X_{mi}} \right) + 1} \tag{15}$$

Writing F for the term in curly brackets and C for the multiplier of that term we have:

$$\begin{aligned} \rho_{mi1} &= \frac{1}{CF + 1} \\ \frac{\partial \rho_{mi1}}{\partial x} &= -\frac{\partial(CF)}{\partial x} \cdot (CF + 1)^{-2} \\ &= -\rho_{mi1}^2 \frac{\partial(CF)}{\partial x} \end{aligned}$$

Our estimates of m , β and v are likely to be the most unreliable of the values we have used

Table A1 Effect of varying certain parameter values on the estimates of mixing matrix elements

	Baseline (see table 5)	(β_m/β_f) times 2	(β_m/β_f) times 0.5	$m_{m2f}/m_{m2f} = 89/4$	$m_{m2f}/m_{m2f} = 55/18$
Men:					
ρ_{m11}	0.96	0.97	0.94	0.96	0.96
ρ_{m12}	0.04	0.03	0.06	0.04	0.04
ρ_{m21}	0.88	0.90	0.88	0.96	0.78
ρ_{m22}	0.12	0.10	0.12	0.04	0.22
ρ_{m31}	0.94	0.95	0.93	0.94	0.94
ρ_{m32}	0.06	0.05	0.07	0.06	0.06
Women:					
ρ_{f11}	0.94	0.95	0.93	0.94	0.94
ρ_{f12}	0.05	0.05	0.05	0.05	0.04
ρ_{f13}	0.01	0.01	0.01	0.01	0.01
ρ_{f21}	0.91	0.70	1.23	0.91	0.91
ρ_{f22}	0.13	0.12	0.14	0.04	0.25
ρ_{f23}	0.02	0.01	0.02	0.02	0.02

to solve equation (15). However, β and v occur in F both in the numerator and the denominator, and as part of a sum. Therefore it is likely that the estimate of ρ is most sensitive to changes in estimates of m .

Errors in the estimation of m may result from small sample size and bias in the identification of contact pairs. Errors will be greater for larger values of ρ .

Table A1 below presents some examples of the effect on ρ of changing the values of selected parameters in the model. Examination of (15) shows that as far as the β 's and v 's are concerned it is only the ratio of $\beta_m:\beta_f$ and $v_m:v_f$ that are important and furthermore, changing either ratio is equivalent to changing the other. We have therefore tabulated only the effects of doubling and halving the ratio $\beta_m:\beta_f$.

We have not tabulated the effect of changing the parameter ratios N_{f2}/N_{f1} , X_{f2}/X_{mi} and Y_{f2}/Y_{mi} . The value of ρ is most sensitive to changing N_{f2}/N_{f1} which is estimated directly from the NSSAL, as is X_{f2}/X_{mi} . The estimate of Y_{f2}/Y_{mi} is based on fairly large numbers of total clinic attenders, and the algebra shows that the mixing matrix estimate will not be particularly sensitive to changing this value.

The most problematic term is m_{mif2}/m_{mif1} . This is estimated from a relatively small number of identified gonorrhoea transmission pairs (table 2), and will exert a significant effect on the mixing matrix estimates. One of the most interesting features of the results which we have presented is the evidence to suggest preferential mixing within stratum two individuals (those with two partners in the last month). Here we focus on ρ_{m2i} . In the results section we calculated the values for ρ_{m2i} using the point estimate of m_{m2f2}/m_{m2f1} equal to 11/72. It seems reasonable to assume that both numerator and denominator are distributed as Poisson random

variables, giving approximate 95% confidence intervals of 4,18 and 55,89 respectively. We have taken a pessimistic estimate of the confidence limits of the fraction by varying both numerator and denominator simultaneously in opposite directions to give broad confidence intervals of 4,89 to 18,55. Comparison with the values expected under proportional mixing assumptions (table 5) show that proportional mixing is on the borderline of credibility with respect to the gonorrhoea transmission pairs data.

As an alternative method of assessing the sensitivity of our solution to changes in the parameters, we have done a Monte Carlo analysis of equation (15), with our uncertainty about the true values of various of the parameters modelled as random variables as follows:

$$\begin{aligned} \beta_m &\sim \text{uniform}(0.5, 0.7) \\ \beta_f &\sim \text{uniform}(0.7, 0.9) \\ v_m &\sim \text{uniform}(0.5, 2.0) \\ v_f &\sim \text{uniform}(0.5, 2.0) \\ Y_{f(1,2)} & Y_{m(1,2,3)} \sim \text{Poisson, with rate estimate from table 1.} \\ m_{(i,j)} &\sim \text{Poisson, with rate estimate from table 2.} \end{aligned}$$

Ninety five per cent confidence intervals are given in table A2. The results show that the mixing estimates are much more stable for the lower activity classes. This is largely because the estimates for ρ_{mij} are based on m_{mif1} and m_{mif2} , and these estimates are based on much smaller numbers in the higher activity classes. Again the values expected under proportional mixing are on the borderline of credibility with respect to the data.

Table A2 Centiles for the mixing matrix from Monte-Carlo sampling

	0.025 centile	0.975 centile
ρ_{m11}	0.9164	0.9803
ρ_{m12}	0.0197	0.0836
ρ_{m21}	0.8159	0.9451
ρ_{m22}	0.0549	0.1841
ρ_{m31}	0.7583	0.9610
ρ_{m32}	0.0390	0.2417
ρ_{f11}	0.8981	0.9607
ρ_{f12}	0.0420	0.0487
ρ_{f13}	0.0097	0.0123
ρ_{f21}	0.4268	1.0000
ρ_{f22}	0.0625	0.2097
ρ_{f23}	0.0111	0.0685

- Renton AM, Whitaker L. Using STD occurrence to monitor AIDS prevention. *Soc Sci Med* 1994;38:1153-65.
- Renton AM, Ison CA, Whitaker L, Kirtland K, Kupek E, Harris JRW. Neisseria gonorrhoeae isolated at St Mary's Hospital, London, 1980-91. *Gemittourin Med* 1993;69:286-9.
- Catchpole MA. Sexually transmitted disease in England and Wales: 1981-1990. *CDR Review* 1992;2:R1-R7.
- Secretary of State for Health. *The health of the nation: a strategy for health in England*. London, HMSO, 1992. (CM 1986).
- Hethcote HW, York JA. Gonorrhoea transmission dynamics and control. In: Castillo-Charez C ed. *Lecture notes in biomathematics*. No 56. New York: Springer Verlag, 1984.
- Whitaker L, Renton AM. A theoretical problem of interpreting the recently reported increase in homosexual gonorrhoea. *Eur J Epidemiol* 1992;8:187-91.
- Garnett GP, Anderson RM. Contact tracing and the estimation of sexual mixing patterns: the epidemiology of gonococcal infections. *Sex Transm Dis* 1993;20:181-91.

- 8 Gupta S, Anderson RM, May RM. Networks of sexual contacts: implications for the pattern of spread of HIV. *AIDS* 1989;3:807-17.
- 9 Anderson RM. Mathematical and statistical studies of the epidemiology of HIV. *AIDS* 1989;3:333-46.
- 10 Jacquez JA, Simon CP, Koopman J, Sattenspiel L, Perry T. Modelling and analysing HIV transmission: the effect of contact patterns. *Mathematical Biosciences* 1988;92:119-99.
- 11 Jacquez JA, Simon CP, Koopman J. Structured mixing: heterogeneous mixing by definition of activity groups. In mathematical and statistical approaches to AIDS epidemiology. In: C Castillo-Chavez ed. *Lecture notes in biomathematics*. Vol 83. Berlin: Springer Verlag, 1989: 301-315.
- 12 Garnett GP, Anderson RM. Factors controlling the spread of HIV in heterosexual communities in developing countries: patterns of mixing between different age and sexual activity classes. *Phil Trans R Soc Lond B* 1993;42:137-159.
- 13 Johnson A, Wadsworth J, Wellings K, Field J. *Sexual attitudes and lifestyles*. Oxford: Blackwell, 1994.
- 14 Rotherberg RB. The geography of gonorrhoea: empirical demonstration of core group transmission. *Am J Epidemiol* 1983;117:688-94.
- 15 Potteratt JJ, Rotherberg RB, Woodhouse DE, Muth JB, Pratts CI, Fogle JS. Gonorrhoea as a social disease. *Sex Transm Dis* 1985;12:25-32.
- 16 Granath F, Gieseke J, Scalia-Tomba G, Ramstedt K, Forssman L. Estimation of a preference matrix for women's choice of male sexual partner according to rate of partner change, using partner notification data. *Mathematical Biosciences* 1991;107:341-48.
- 17 Joyce CF, Garnett GP, Renton AM, Whitaker L, Ison CA. Persistence of multiple serovars of *Neisseria gonorrhoeae*. *Sex Transm Dis* 1993;20:105-9.
- 18 Ison CA, Whitaker L, Renton A. Concordance of auxotype/serovar classes of *Neisseria gonorrhoeae* between sexual contacts. *Epidemiol Infect* 1992;109:265-71.
- 20 Swinton J, Garnett GP, Brunham RC, Anderson RM. Gonococcal infection, infertility and population growth: 1. Endemic states in behaviourally homogeneous populations. *IMA J Math Appl Med Biol* 1992;9:107-26.