ORIGINAL ARTICLE

Modelling the effect of *Chlamydia trachomatis* testing on the prevalence of infection in England: what impact can we expect from the National Chlamydia Screening Programme?

Matthew David Dorey, Yoon Hong Choi, Kate Soldan, Emilia Vynnycky

NOTICE OF RETRACTION

The authors and the editors are issuing a retraction of the article entitled "*Modelling the effect of Chlamydia trachomatis testing on the prevalence of infection in England: what impact can we expect from the National Chlamydia Screening Programme?*", which was published online in *Sexually Transmitted Infections* on June 26, 2012 (DOI:10.1136/sextrans-2011-050126).

The authors noticed a coding error as a result of ongoing work using the model described in the article. They are working on a revised model and intend to publish a corrected version of the article when it is finalised. The editors of the journal, in agreement with the authors, have decided that the original version of the article should therefore be retracted.

Modelling the effect of *Chlamydia trachomatis* testing on the prevalence of infection in England: what impact can we expect from the National Chlamydia Screening Programme?

Matthew Dorey[1,2], Yoon Hong Choi[3], Kate Soldan[1], Emilia Vynnycky[2]

1 HIV/STI Department, Health Protection Agency, Colindale, London, NW9 5EQ.

2 Modelling and Economics Unit, Health Protection Agency, Colindale, London, NW9 5EQ.

3 Immunisation, Hepatitis and Blood Safety Department, Health Protection Agency, Colindale, London, NW9 5EQ.

Appendix A: Additional modelling material

Further information on the partner change rates

We use the partner change rates estimated from NATSAL as presented in Choi et al [1]. For each of the age bands above the NATSAL age range (45-54, 55-64 and 65-77 years), we assume that the partner change rate is 50% of the previous one. Choi et al [1] found that using successive reductions of 30% and 70% led to no significant difference in their results, so the value of 50% is adopted throughout.

The age of sexual debut is assumed to be 14 years. We define the partner change rate for the 14-15 year old age group in the same way as in Choi et al [1]: the NATSAL estimate for 16-19 year olds multiplied by the proportion of NATSAL respondents who reported 14 or 15 years as their age of sexual debut.

	Range for number of new partners per year (min max) / Activity group					
	Low		**Medium**		**High**	
Age (years)	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
16-19	0-3	0-2	4-7	3-5	8+	6+
20-24	0-3	0-1	4-9	2-3	10+	4+
25-29	0-2	0-1	3-5	2-3	6+	4+
30-34	0-1	0-1	2-3	2-3	4+	4+
35-44	0-1	0-1	2-3	2-3	4+	4+

Table A1: *Ranges for the number of new partners per year associated with each activity group (min-max), calculated so that the low, medium and high activity groups comprise 80%, 15% and 5% respectively of the population. These values were calculated from the responses to the NATSAL survey by individuals who provided a sample for chlamydia testing. For age groups in the model outside of those considered in the NATSAL, we used the assumptions of Choi et al* [1].

Model equations

The number of susceptible, treated and infected individuals of sex k, age m months and sexual activity group r at a given time are given by the following equations.

$$\begin{aligned}
S_{k,m+1,r} &= S_{k,m,r} - \lambda_{k,m,r}S_{k,m,r} + \frac{1}{D}I_{k,m,r} + \rho T_{k,m,r} \\
&\quad - \mu_{k,m}\left(S_{k,m,r} + \rho T_{k,m,r}\right) \\
T_{k,m+1,r} &= \phi_{k,m,r}\lambda_{k,m,r}S_{k,m,r} + \tau^{S}_{k,m,r} + \tau^{T}_{k,m,r} \\
I_{k,m+1,r} &= I_{k,m,r} + \left(1-\phi_{k,m,r}\right)\lambda_{k,m,r}S_{k,m,r} + (1-\rho)T_{k,m,r} - \frac{1}{D}I_{k,m,r} \\
&\quad - \tau^{S}_{k,m,r} - \tau^{T}_{k,m,r} - \mu_{k,m}(I_{k,m,r} + (1-\rho)T_{k,m,r})
\end{aligned} \tag{1}$$

The model uses a time step of one month and ageing occurs concurrently with other events in the model, so time dependence is not made explicit in the model equations.

The force of infection at a given time differs across the sex, age and activity groups of the model. Its expression is identical to that used in Choi et al [1], and for convenience, has been replicated here for individuals of sex k, age group g and activity group r:

$$\lambda_{k,g,r} = \sum_{h=0}^{9}\sum_{s=0}^{2} C_{1-k,g,h,r,s}M_{1-k,g,h,r,s}\zeta_{1-k,r,s}\frac{I_{1-k,h,s}}{n_{1-k,h,s}}. \tag{2}$$

Our model additionally includes screening and the notification of partners of index cases identified by screening. To attain the required level of positivity among those tested, the number of infected individuals that needed to be screened was first calculated using the screening rate σ, the total number of infected individuals in each screening age group and the positivity-to-prevalence ratio. Infected individuals identified by testing are allocated to activity groups proportional to the activity group composition of the entire infected population.

For each sex and activity group, the infected individuals found by screening in each age group (i.e. 16-19 & 20-24) are assumed to be evenly distributed by age. This is a simplification as, for example, older males (20-24) are more likely to be tested in GUM settings than younger ones (16-19). A proportion χ of these individuals will submit contact details for tracing of partners (see below) while the remaining proportion $(1-\chi)$ will be at risk of re-infection by untreated partners. The proportion of these latter individuals that are re-infected is denoted by ϖ.

As a result, the total number of individuals that are treated after being tested ($\tau^{S}_{k,m,r}$) is given by the sum of a) the number of individuals who are screened and who provide details of their contacts ($\chi I^{X}_{k,m,r}$) and b) the number of individuals who are screened that did not provide details of their contacts but also managed to avoid re-infection ($(1-\varpi)(1-\chi)I^{X}_{k,m,r}$), i.e.:

$$\tau^{S}_{k,m,r} = \chi I^{X}_{k,m,r} + (1-\varpi)(1-\chi)I^{X}_{k,m,r}. \quad (3)$$

The individuals who are screened and are re-infected are assumed to remain in the infected class.

The number of additional infected individuals traced from the partners of screened infected individuals, who are subsequently treated ($\tau^{T}_{k,m,r}$) is given by the following expression:

$$\tau^{T}_{k,m,r} \approx \vartheta^{X}_{k,m,r} I_{k,m,r}, \quad (4)$$

where $\vartheta^{X}_{k,m,r}$ approximates the proportion of infected individuals of sex k, monthly age group m and activity group r that are notified because of testing. Equation (4) uses the same partner change rates and mixing patterns as equation (2) and substitutes the number of index cases that supply contacts for the number of infected individuals in that expression, as follows:

$$\vartheta^{X}_{k,g,r} = \sum_{h=0}^{9}\sum_{s=0}^{2} C_{1-k,g,h,r,s} M_{1-k,g,h,r,s} \frac{\chi I^{X}_{1-k,h,s}}{N_{1-k,h,s}}. \quad (5)$$

The expression (5) is calculated across age groups and we assume that these can be evenly distributed across the monthly cohorts that comprise each age group.

Table A2: *Summary of the definitions of the variables used in the model.*

Variable	Description
$S_{k,m,r}$	Number of susceptible individuals of sex k, age m and activity group r at a given time.
$T_{k,m,r}$	Number of individuals of sex k, age m and activity group r currently undergoing treatment at a given time.
$\lambda_{k,m,r}$	The force of infection on individuals of sex k, age m months, activity group r.
$I_{k,m,r}$	Number of asymptomatically infected individuals of sex k, age m and activity group r at a given time.
$I^{X}_{k,m,r}$	Number of asymptomatically infected individuals found by testing of sex k, age m and activity group r at a given time.
$\tau^{S}_{k,m,r}$	Number of infected individuals of sex k, age m and activity group r that are treated as a result of asymptomatic testing at a given time.
$\tau^{T}_{k,m,r}$	Number of infected individuals of sex k, age m and activity group r that are treated at a given time as a traced partner of a screened individual.
$N_{k,g,r}$	The total number of individuals of sex k in age group g and activity group r.

The transmission rates

Here we provide a more general explanation of our construction of transmission rates than that given in the main text. We define β to be the transmission rate per partnership in a partnership consisting of two individuals in the low activity group. We use this as the basis for transmission rates per partnership involving at least one individual from outside the low activity group.

The transmission rate per partnership for a partnership consisting of two individuals in the medium activity group is defined to be $\alpha\beta$, where α is a scaling factor between zero and one that reflects the fact that fewer sex acts occur in partnerships between individuals with higher activity levels; and β is defined as above. The transmission rate for a partnership involving two high activity individuals is defined to be $\alpha^2\beta$.

Table A3: *Summary of the parameters used in the model.*

Parameter	Description	Assumed value
$\mu_{k,m}$	Mortality rate (sex k, age m).	Assumed to be identical to that for 2006 in England [2].
D	The average duration of asymptomatic infection.	12 months [3,4]. Additional scenario of 6 months also run.
ρ	Treatment efficacy.	95% (Data from NCSP, also scenarios of 90% and total efficacy were considered)
$\phi_{k,m,r}$	The proportion of newly infected individuals of sex k, age m and activity group r that seeks treatment.	Estimated by fitting model predictions to the observed prevalence data from the NATSAL survey. Assumed the same for all age, sex and activity groups.
χ	The proportion of tested individuals found to be infected that have all of their contacts traced	In the base case, we assume a value of 0.2. Alternative value of 0.4 is also considered.
ϖ	The proportion of treated individuals who do not provide contact information and are subsequently re-infected by an untreated partner.	In the base case, we assume a value of 0.4. Alternative values of 0.2 and 0.6 are considered.
β	Transmission rate associated with a partnership between two individuals in the low activity group.	Estimated by fitting model predictions to the observed prevalence data from the NATSAL survey.
α	Scaling factor applied to β to describe transmission rates for partnership involving at least one partner from medium or high activity group	Estimated by fitting model predictions to the observed prevalence data from the NATSAL survey.
$\zeta_{k,r,s}$	Transmission rate between an individual of sex k, activity group r, and individuals of sex 1-k, activity group s	Composed from the transmission rate β for partnerships between low activity individuals and the scaling factor α, according to Table A4.
$C_{k,g,h,r,s}$	The rate of change of the number of partnerships between individuals of sex k, age group g and activity group r, and individuals of sex 1-k, age group h and activity group s	Assumed to be identical to that used in [1]. $1/C_{k,g,h,r,s}$ reflects the average duration of a partnership between individuals of sex k, age group g and activity group r, and individuals of sex 1-k, age group h and activity group s.
$M_{k,g,h,r,s}$	Sexual mixing between individuals of sex k, age group g and activity	Assumed to be identical to that used in [1].

	group r, and individuals of sex 1-k, age group h and activity group s	

Table A4: *Expressions for the transmission rates per partnership associated with partnerships between persons in different activity groups. In a) the transmission rate associated with the person from the lowest activity group in the partnership dominates the transmission rate, in b) the transmission rate associated with the person from the highest activity group in the partnership dominates the transmission rate and in c) the transmission rate among the persons in the lowest activity group in the partnership dominates the transmission rate and the transmission in a partnership comprising two high activity individuals is the same as one that involves at least one medium activity individual.*

a) Lowest activity individuals in the partnership dominate

	High	**Medium**	**Low**
High	$\alpha^2\beta$	$\alpha\beta$	β
Medium	$\alpha\beta$	$\alpha\beta$	β
Low	β	β	β

b) Highest activity individuals in the partnership dominate

	High	**Medium**	**Low**
High	$\alpha^2\beta$	$\alpha^2\beta$	$\alpha^2\beta$
Medium	$\alpha^2\beta$	$\alpha\beta$	$\alpha\beta$
Low	$\alpha^2\beta$	$\alpha\beta$	β

c) Alternative scenario

	High	**Medium**	**Low**
High	$\alpha\beta$	$\alpha\beta$	β
Medium	$\alpha\beta$	$\alpha\beta$	β
Low	β	β	β

For partnerships involving individuals from two different activity groups, we consider three different cases: either the transmission rate associated with the lowest activity individual dominates the transmission rate per partnership (as considered in the main text) or that associated with the highest activity individual dominates. The effects of these assumptions on the expressions for the transmission rates per partnership in terms of α and β are shown in Table A4. In estimating α and β, we considered each of these cases separately and found that only the case in which the transmission rate associated with low activity individuals dominates leads to viable and best-fitting estimates for all the parameters.

Fitting the model to data

The equation for the log-likelihood deviance used in the fitting process is as follows:

$$Dev = -2\sum_{i=0}^{36}\Bigg(P_i \log \mathrm{y_i} + (N_i - P_i)\log(1 - y_i) - P_i \log\left({}^{P_i}/_{N_i}\right) - (N_i - P_i)\log(1 - {}^{P_i}/_{N_i})\Bigg), \quad (6)$$

where P_i and N_i are respectively the observed number positive and the observed number tested for the i^{th} data point of the NATSAL prevalence data, and y_i is the model prediction for the proportion of infected individuals for the i^{th} data point. The deviance was minimised using the Brent Method [5].

The data points were drawn from the NATSAL prevalence data with a weighting applied to account for biases present in responses by age, gender and ethnicity [6]. The prevalence in each activity group was calculated by stratifying all test results for chlamydia according to the number of new partners in the previous year. For each age group, the denominators were calculated as 5%, 15% and 80% of the weighted populations providing a chlamydia test result and the number of positives was then approximated from the positive tests by working backwards through the number of positives per number of new partners in the last year.

The data points and the model outcomes for the best fitting parameter values are shown together in Figure A1. While the overall model prevalence at equilibrium of 4.5% in 16-24 year olds is greater than that observed in NATSAL (2.7% for men and 3.0% for women aged 18-24 years), it is still within the 95% confidence intervals for each sex (1.2% to 5.8% for men and 1.7% to 5.0% for women) [7].

The likely reason for our model producing a higher prevalence than NATSAL can be seen in Figure A1, where the model predicts non-zero prevalence for 16-19 year olds in the highest activity groups while there no corresponding infections in the data. Though these infections are in 5% of the population, they contribute significantly to overall prevalence in the model prediction.

To investigate what parameter values led to prevalence closer to that seen in NATSAL, we ran the model to equilibrium on multiple randomly generated parameter

sets until 100 achieved an equilibrium in the range of 2.5% to 2.9% in men aged 18-24 years and 2.8% to 3.2% for women aged 18-25 years. None of these sets of parameters contained values that were admissible. Therefore choosing parameters that led to prevalence at the high end of the NATSAL was the preferable option.

Converting between measures of partner notification

In our model we measure the extent of partner notification as the proportion of those testing positive that have all their contacts successfully traced. In practice, partner notification is usually characterised in terms of the average number of partners contacted per index case. Because our model is population based, it was convenient to formulate the partner notification process in terms of known quantities at the population level. It is possible to convert between these two types of partner notification measure, using values for the median number of new partners per year from Table A1 and the following equation:

$$PPI \approx \frac{\chi \sum_{g=3}^{4} \sum_{k=0}^{1} \sum_{r=0}^{2} z_{k,g,r} \Gamma_{k,g,r} P_{k,g,r}}{\sum_{g=3}^{4} \sum_{k=0}^{1} \sum_{r=0}^{2} z_{k,g,r} \Gamma_{k,g,r}}, \quad (7)$$

where PPI is the average number of partners that are contacted per index case; χ is the proportion of individuals testing positive that have all of their contacts traced; $\Gamma_{k,g,r}$is the proportion of eligible adults (i.e. those aged 16-24 years) who are in a given age group g, sex k and activity class r; $P_{k,g,r}$ is the median number of new partners in the previous year for an individual of sex k, age g, and activity group r as calculated from Table A1; and $z_{k,g,r}$ is the proportion of individuals of sex k in age group g and sexual activity group r that are infected. Table A5 gives some examples of conversions between the proportion of index cases with all contacts traced and the estimated number of contacts traced per index case when the proportion infected is similar to that at the start, i.e. when the infection prevalence is assumed to be at equilibrium.

Table A5: *Example values from the approximate conversion between the formulation of partner notification in the model and that used in screening programme, calculated using equation (7).*

Percentage of persons testing positive who have all their contacts traced	Number of partners tested per person testing positive
10%	0.26
20%	0.53
30%	0.79
40%	1.05
50%	1.32
60%	1.58

Although this is a quantitative comparison, the two measures do lead to qualitatively different populations being identified by partner notification. For example, because the prevalence of infection is highest in the high activity group, our model will identify a larger number of individuals in the highest activity group than in the other activity groups, although in reality these individuals are engaged in short term partnerships that are least likely to be identified. Moreover, our analyses probably underestimate the impact of notifying partners since we calculate the number of partners that are notified using the number of new partnerships that individuals form each year in each activity group rather than the numbers of *recent* partnerships. Our approach means that the proportion of the population that our model targets are perhaps less likely to include infected individuals than are models which implement partner notification using the number of recent partnerships. The latter approach is consistent with the way partner notification is implemented in reality, i.e. persons found to be positive are asked to provide details of their recent partners rather than partners that they had during the previous year.

Effects of PN and re-infection in the model

Figure A2 shows the effect of partner notification and re-infection on the decrease in prevalence. In Figure A2a, the model prediction for the proportion of infected individuals aged 16-24 years is compared for different levels of partner notification. The proportion of individuals who do not provide contacts and are then re-infected by an untreated partner is fixed throughout at 40%. If the partner notification is

increased so that all partners of 60% of individuals are contacted then the model predicts that the median prevalence in 2009/10 would differ by 0.2% from that predicted using the base case assumptions (where all partners of 20% of persons are contacted).

Figure A2b shows the effect of different levels of re-infection on the effects of testing. The base case scenario is for 40% of individuals who do not provide contact details to be re-infected by an untreated partner. If this is set at 60% (so that more re-infections occur compared to the base case), the model predicts that median prevalence will be 0.2% higher in 2009/10 than predicted in the base case. Alternatively, if re-infection is set at 20% (so that fewer re-infections occur), the model predicts that median prevalence will be 0.2% lower in 2009/10 than predicted in the base case.

Other models of partner notification

Partner notification has been implemented in other deterministic models. In Armbruster et al [8], in the context of a general endemic disease, a mechanism to describe contact tracing is developed, which is quite complex though mathematically tractable; however, the population considered is not stratified in any way. Eames and Keeling [9] included contact tracing in a deterministic pair-wise equation model, demonstrating the need for inclusion of individual level events within such a model. Once again the population is not stratified; to stratify the population as we have done would lead to a complicated model that may as well be individual-based.

Additional scenarios considered

As shown in Table 2 (main text), additional scenarios were considered during the parameter estimation process. These included considering a duration of asymptomatic infection of six months, in addition to various combinations of the transmission rate per partnership for partnerships involving different activity groups (see Table A4), the assortativity of mixing by risk group and the treatment efficacy.

Whilst in most scenarios the fitting routine was able to converge upon a solution that minimised the log-likelihood deviance between model predictions and data, in some cases the parameter estimates had to be rejected due to the unrealistic model predictions being produced. This usually included an equilibrium in which the

prevalence was significantly lower in the high activity group when compared to that of the low activity group. In general, assumptions that the parameter reflecting assortativity of mixing was 0% or 90% (corresponding to the assumption that persons in differing activity groups have little contact or contact each other almost exclusively respectively) led to implausible values for the prevalence of infection.

We generated additional rejection samples for the best fitting parameter values for six months duration of asymptomatic infection. The result of model runs based on these rejection samples is given in Figure A3. The median equilibrium prevalence is predicted to be 5.2% (IQR: 4.6%-5.8%), decreasing to 4.2% (IQR: 3.8% to 4.8%) by 2010/11. This is a reduction of 19% relative to baseline. The projections of sustained and increased testing are respectively predicted to produce reductions of 2.4% and 3.6% per year relative to 2010/11.

References

1. Choi YH, Jit M, Gay N, et al. Transmission dynamic modelling of the impact of human papillomavirus vaccination in the United Kingdom. Vaccine. 2010;28(24):4091-4102.

2. Office for National Statistics. Population estimates and mortality statistics. Deaths registered in 2006 and population estimates for years 2000 to 2010.

3. Turner KME, Adams EJ, Lamontagne DS, et al. Modelling the effectiveness of chlamydia screening in England. Sexually transmitted infections. 2006;82(6):496-502.

4. Golden M, Schillinger J, Markowitz L, et al. Duration of Untreated Genital Infections With Chlamydia trachomatis: A review of the literature. Sexually Transmitted Diseases. 2000;27(6):329-337.

5. Press WH, Teukolsky SA, Vetterling WT, et al. Numerical Recipes in C++, Third Edition. 3rd ed. Cambridge University Press; 2007.

6. McCadden A, Fenton KA, McManus S, et al. Chlamydia trachomatis Testing in the Second British National Survey of Sexual Attitudes and Lifestyles: Respondent Uptake and Treatment Outcomes. Sexually Transmitted Diseases. 2005;32(6):387-394.

7. Fenton KA, Korovessis C, Johnson AM, et al. Sexual behaviour in Britain: reported sexually transmitted infections and prevalent genital infection. The Lancet. 2001;358(9296):1851-1854.

8. Armbruster B, Brandeau ML. Optimal mix of screening and contact tracing for endemic diseases. Mathematical biosciences. 2007 ;209(2):386-402.Available from: http://www.ncbi.nlm.nih.gov/pubmed/17428503

9. Eames KTD, Keeling MJ. Contact tracing and disease control. Proc. R. Soc. Lond. B. 2003 ;2702565-2571.

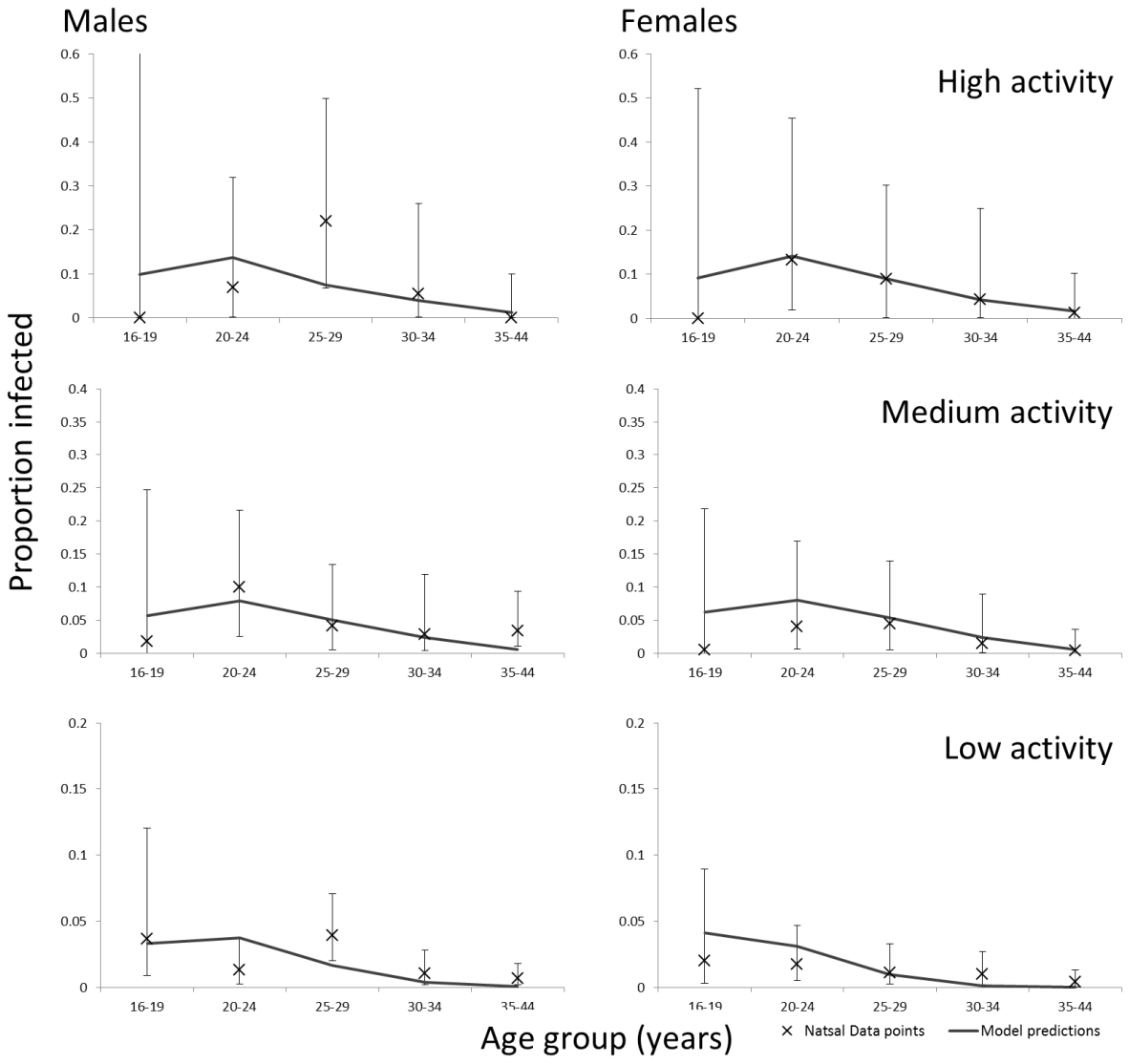

Figure A1: *Comparison between data on the age-specific proportion of males and females in the high, medium and low activity groups found to be positive during the NATSAL survey and the best fitting model predictions. Ranges on the data points show 95% confidence intervals.*

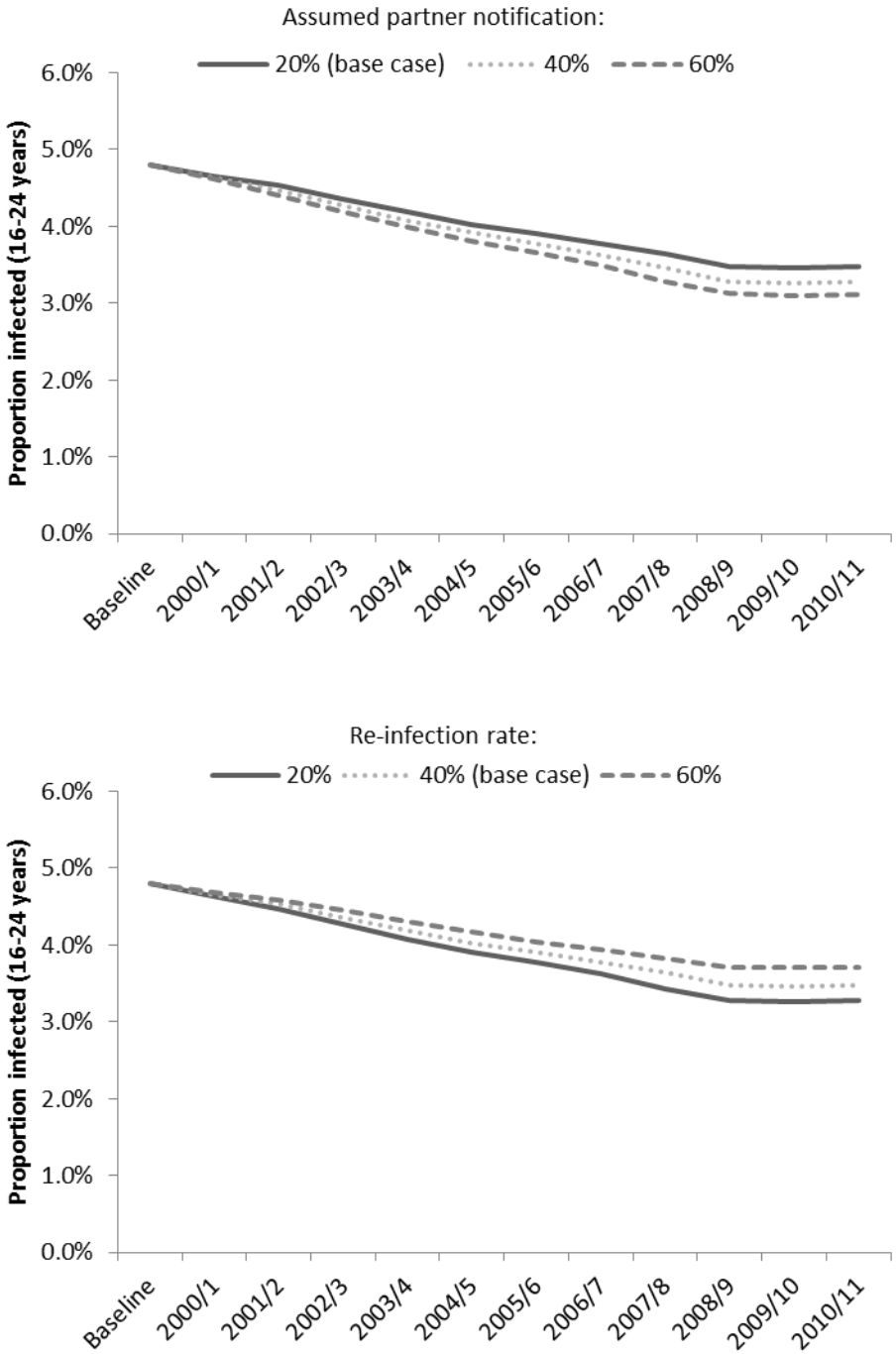

Figure A2: *The effect of altering the levels of partner notification and re-infection on the impact of testing. Both subfigures show the proportion of 16-24 year olds infected (using the median of 100 outcomes). In a) 40% of treated individuals that did not provide contacts were re-infected by an untreated partner and the percentage of partners of treated individuals who were notified was increased between 20% and 60%. In b) the percentage of partners of treated individuals notified was fixed at the*

base case value of 20%, while the re-infection rate for individuals that did not provide contacts was increased between 20% and 60%.

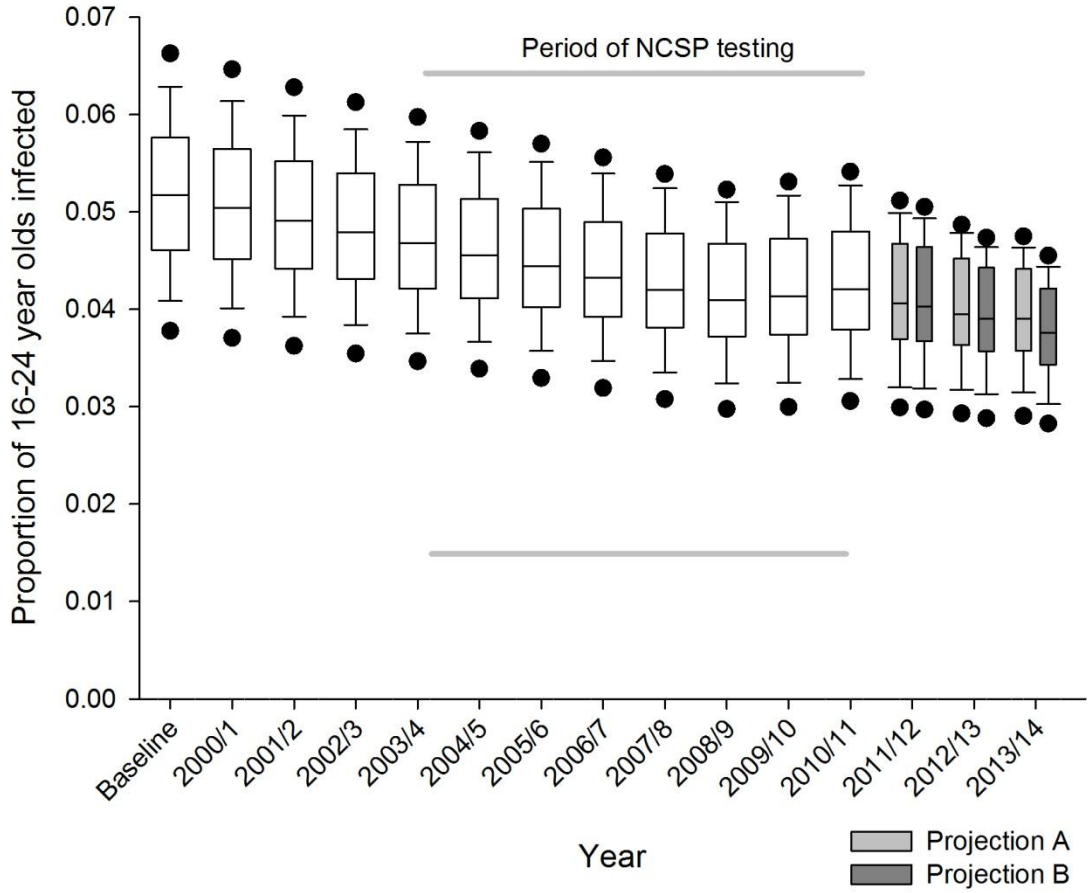

Figure A3: *Similar to Figure 2, showing model predictions obtained assuming that the duration of asymptomatic infection is six months. One hundred parameter sets were chosen using rejection sampling based on the best fitting parameter estimates shown in Table 4. The testing coverage and positivity are as presented in Table 2, while the assumptions for projections A and B are presented in Table 3.*

Modelling the effect of *Chlamydia trachomatis* testing on the prevalence of infection in England: what impact can we expect from the National Chlamydia Screening Programme?

Matthew Dorey[1,2], Yoon Hong Choi[3], Kate Soldan[1], Emilia Vynnycky[2]

1 HIV/STI Department, Health Protection Agency, Colindale, London, NW9 5EQ.

2 Modelling and Economics Unit, Health Protection Agency, Colindale, London, NW9 5EQ.

3 Immunisation, Hepatitis and Blood Safety Department, Health Protection Agency, Colindale, London, NW9 5EQ.

Appendix B: Additional information concerning testing and positivity data.

Figures for the number of chlamydia tests performed and the number of positive diagnoses made by the NCSP were taken from the NCSP data set [1]. These data are considered to be a complete collection of all tests and positive diagnoses made by the screening programme.

Figures for the number of tests and the number of positive diagnoses made in GUM settings were taken from the KC60 data for the years 2003/4 to 2008/9 [2], which includes all sexual health screens performed in GUM clinics in England that involved a test for chlamydia.

For the years 2000/1 to 2002/3, data on the number of positive tests were available but not for the number of tests, so the number of tests for these years was imputed using linear regression. While this is a very simple model for the extrapolation, the small number of data points required made this an appropriate choice. We considered two alternative scenarios (results not shown) for the imputed values: holding positivity constant at 2003/4 levels and decreasing between 2000/1 and 2002/3 at half the rate obtained by linear regression for the years 2003/4 to 2009/10. These scenarios did not significantly alter model predictions.

Figures for the tests and diagnoses made in Non-NCSP / Non-GUM settings from the year 2008/9 onward were taken from the GUMCAD data [2]; data from before

these years were not collected as part of this data set. Figures for the mid-year population estimates were taken from ONS data [3].

The infection prevalence is assumed to be at equilibrium in 2000 and we impute the missing values for testing and positivity for the years 2000/1 to 2002/3 for GUM and for the years 2000/1 to 2007/8 for Non-NCSP/Non-GUM settings. For GUM settings we performed linear regression on both the numbers of persons tested for infection and extrapolated accordingly. The extrapolated figures for the number of individuals tested were then divided by the population estimate for that year to give the estimates of coverage given in Table B1. Considering non-NCSP/non-GUM settings the numbers of males tested or found to be positive suggested that coverage was low and that positivity was high, which was probably indicative of symptomatic individuals seeking testing. As treatment seeking behaviour is not varied over time in the model, we assumed that for males the number of tests for chlamydia carried out in non-GUM/non-NCSP settings remained constant, along with the number of positive diagnoses. For females, we noted a trend similar to that of GUM, so we extrapolated according to that trend.

For each year, we have assumed that each test and positive diagnosis in each setting corresponds to a single individual. This means that we may have overestimated of the number of tests and diagnoses as some individuals may have been tested more than once. We also consider all tests to be of asymptomatic individuals and that none of the tests include people who were tested as the partner of someone previously tested.

GUM figures for chlamydia positives have been derived from S1 and S2 codes in the KC60 data. These are for STI screens with and without an HIV test. It is possible that some individuals tested for chlamydia who were not tested for gonorrhoea would be missed from these values as both tests have to be performed in order to be coded. This means that the number of persons tested in GUM could be underestimated.

For simplicity, we have not taken into consideration the geographic heterogeneity of those tested. For example, we consider the additional tests carried out by the NCSP from 2003/4 to have been part of national coverage, whereas in reality these tests were localised within those areas that first introduced additional testing.

Table B1 shows the percentage of 16-24 years olds tested by setting and by year. Table B2 shows the percentage of those tested who were found to be positive (subject to the assumptions above) by setting and by year. Table B3 provides the number of tests performed. In each table, imputed values are indicated with grey cells.

References

1. Data from the National Chlamydia Screening Programme. Available from: http://www.chlamydiascreening.nhs.uk/ps/data/data_tables.html

2. Data from the Genitourinary Medicine Clinic Activity Dataset (GUMCAD).

3. Office for National Statistics. Population estimates and mortality statistics. Deaths registered in 2006 and population estimates for years 2000 to 2010.

Table B1: *Percentage of all 16-24 year olds tested per year for chlamydia, by setting. Grey shaded cells contain imputed values using the methods detailed in Appendix B. The population estimates that are the denominators for these figures are given in Table 2.*

	GUM		NCSP		Non-NCSP/Non-GUM	
Year	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
2000/1	2.4	3.4	No NCSP testing		0.9	1.4
2001/2	3.1	4.4			0.9	2.5
2002/3	3.7	5.3			0.9	3.5
2003/4	4.3	6.4	0.1	0.6	0.8	4.4
2004/5	4.8	7.1	0.3	2	0.8	5.4
2005/6	5.3	7.7	0.7	3.1	0.8	6.2
2006/7	5.7	8.2	1.1	4.1	0.8	7.1
2007/8	6.4	9.5	3	7.4	0.8	8
2008/9	6.9	10.5	7.7	15.8	0.7	8.8
2009/10	6.9	10.4	13.8	23	1.0	9.1
2010/11	6.1	10.0	16.7	25.7	1.1	9.9

Table B2: *Summary of the assumed percentages of 16-24 year olds who attended GUM, NCSP and non-NCSP/non-GUM clinics who tested positive per year for chlamydia. Grey shaded cells contain imputed values using the methods detailed in Appendix B. Denominators are given in Table B3.*

	GUM		NCSP		Non-NCSP/Non-GUM	
Year	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
2000/1	19.7	27	No NCSP testing		12.7	10.5
2001/2	17.5	22.9			12.7	9
2002/3	17.3	21.5			12.7	8.4
2003/4	16.8	19.3	18.7	10.9	12.7	8.1
2004/5	16.5	18.3	17	11.5	12.7	7.9
2005/6	15.8	16.6	14.1	10.6	12.7	7.7
2006/7	15.5	15.1	12.8	10.6	12.7	7.6
2007/8	14.7	13.7	9.1	9.7	12.7	7.5
2008/9	13.7	12.6	7.1	7.8	12.7	7.5
2009/10	12.7	11.5	5.1	6.4	11.1	6.3
2010/11	12.2	10.7	4.4	5.7	9.7	5.6

Table B3: *Summary of the number of chlamydia tests in 16-24 year olds, carried out in GUM, NCSP and non-NCSP/non-GUM settings. Grey cells contain imputed figures.*

	GUM		NCSP		Non-NCSP/Non-GUM	
Year	**Males**	**Females**	**Males**	**Females**	**Males**	**Females**
2000/1	64,187	88,756	No NCSP testing		23,761	37,707
2001/2	83,567	116,371			23,761	66,208
2002/3	102,947	143,987			23,761	94,709
2003/4	122,957	177,025	1,347	16,183	23,761	123,210
2004/5	143,538	200,574	8,471	55,378	23,761	151,711
2005/6	160,134	223,991	20,882	89,911	23,761	180,212
2006/7	175,848	241,454	34,917	120,341	23,761	208,713
2007/8	201,455	284,034	94,073	221,669	23,761	237,214
2008/9	220,723	316,764	244,794	477,274	23,761	265,715
2009/10	220,675	315,841	441,861	699,894	31,522	277,353
2010/11	197,301	302, 550	536,786	781,975	35,287	299,327