**Discrepancies between observed data and predictions from mathematical modelling of the impact of screening interventions on *Chlamydia trachomatis* prevalence**

This project contains the code for a model to investigate two hypotheses about factors that might attenuate the effects of a chlamydia screening intervention: the existence of long-lasting partial immunity; and differential chlamydia test coverage according to the risk of being infected. The model uses data about sexual behaviour and the prevalence of *C. trachomatis* in the general population of Great Britain from Natsal-2 and Natsal-3 and time-series data about chlamydia testing and diagnoses in England and across the same time period. The model is written in the R language (currently version 3.4.0). It also produces code in C++ language to compute the differentials of a system of differential equations (which are solved in R again).

Original Model Developer: Joost Smid  
Email: [joost.smid@ispm.unibe.ch](mailto:joost.smid@ispm.unibe.ch)  
Affiliations: Institute of Social and Preventive Medicine (ISPM), University of Bern

**Code organization**

All the model scripts, input data and outputs are stored in the main directory and 3 sub-directories. The main directory also contains files to run model variants simultaneously on a computing cluster. These model variants are the four main models, which differ because they contain different fixed parameters, and 24 additional variants of these models, run as sensitivity analyses. Separate models can also be ran one-by-one without using a computing cluster.

*Main directory files*

The following two R scripts run the model using the input files from the data/ directory and the model code from the modelfiles/ directory. They generate the posterior parameter distributions and figures which are stored in the results/ and figures/ directories, respectively. Instructions for running the code are provided in the next section.

run\_multicore.R: Contains the main algorithm calling other routines in modelfiles/ to read in the data files in data/ and runs the MCMC fitting routine. For each model type, five MCMC chains are run. Results for each model (28 different) and each MCMC run per model (5 different) are stored in the results/ directory.

analyze\_mcmc.R: Uses the results for each model stored in the results/ directory to produce the summary results, Figures and Tables. The MCMC chains for one model are combined. Figures and Tables are stored in the figures/ directory.

*modelfiles*

Contains all the specific R functions used in the analysis and by the main directory scripts.

*data*

Contains all the input data and model parameters specifying sexual behavior and C. trachomatis prevalence used by the R scripts. Parameters used in the model to specify sexual behavior and chlamydia prevalence are stored as .Rda files and were obtained from the second and third British National Survey of Sexual Attitudes (Natsal). 1,2 Further, we used data from Chandra et al (Eurosurveillance, 2017)3 providing maximum and minimum estimates for the number of tests and diagnoses in England from 2000-2012.

*results*

MCMC trace results for each model (28 different) and each MCMC run per model (5 different) are stored in the results/ directory as trace*x*.RData files, where *x* is an integer. Every five consecutive .RData files (e.g. trace1.RData – trace5.RData) are results of the same model, but with different MCMC starting conditions.

*Figures*

Summary results, Figures and Tables for each model variant (28 in total) are stored in the figures/ directory.

**Instructions for running the model**

To run the model and analyze its results perform the following steps:

Source run\_multicore.R to generate MCMC traces. To run one model separately, lines 1-3 in this file should be replaced by “args <- c(x,y)”, where x is a number from 1 to 28 (one of the model types), and y is a number between 1 and 5 (one MCMC chain). This produces the MCMC trace and places it in the results/ directory.

Alternatively, one can run all models simultaneously as a batch job on a computing cluster using the files “run\_final.job” and the parameter combinations “parameters.txt” (specifying combinations of model types and MCMC runs), for each job.

Source analyze\_mcmc.R to produce the summary results, Figures and Tables. To produce these for one model separately, lines 1-3 in this file should be replaced by “args <- x”, where x is a number from 1 to 28 (one of the model types). This produces summary results, Figures and Tables for one model and places it in the figures/ directory.

Alternatively, one can analyze all models simultaneously as a batch job on a computing cluster using the files “run\_final\_analyze.job” and the parameter combinations “parameters\_analyze.txt” (specifying the model types), for each job.

1 Fenton, K. A. *et al.* Sexual behaviour in Britain: reported sexually transmitted infections and prevalent genital Chlamydia trachomatis infection. *Lancet* **358**, 1851-1854, doi:10.1016/S0140-6736(01)06886-6 (2001).

2 Sonnenberg, P. *et al.* Prevalence, risk factors, and uptake of interventions for sexually transmitted infections in Britain: findings from the National Surveys of Sexual Attitudes and Lifestyles (Natsal). *Lancet* **382**, 1795-1806, doi:10.1016/S0140-6736(13)61947-9 (2013).

3 Chandra, N. L. *et al.* Filling in the gaps: estimating numbers of chlamydia tests and diagnoses by age group and sex before and during the implementation of the English National Screening Programme, 2000 to 2012. *Euro Surveill* **22**, doi:10.2807/1560-7917.ES.2017.22.5.30453 (2017).