Sexual Health, 2019, **16**, 514–522
doi:10.1071/SH18235_AC © CSIRO 2019

Supplementary Material

Agent-based modelling study of antimicrobial-resistant *Neisseria gonorrhoeae* transmission in men who have sex with men: towards individualised diagnosis and treatment

Adam K. Zienkiewicz[A,B,*], *Nicolás Verschueren van Rees*[A,B,*], *Martin Homer*[A], *Jason J. Ong*[C,D,E], *Hannah Christensen*[E], *Darryl Hill*[F], *Katharine J. Looker*[E], *Paddy Horner*[E], *Gwenda Hughes*[G,H] *and Katy M. E. Turner*[B,E,I]

[A]Department of Engineering Mathematics, University of Bristol, Bristol BS8 1UB, UK.

[B]School of Veterinary Sciences, University of Bristol, Langford House, Langford, Bristol BS40 5DU, UK.

[C]Clinical Research and Development, Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK.

[D]Central Clinical School, Monash University, Clayton, Vic. 3800, Australia.

[E]Population Health Sciences, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK.

[F]School of Cellular and Molecular Medicine, University of Bristol, Biomedical Sciences Building, University Walk, Bristol BS8 1TD, UK.

[G]Instituto de Medicina Tropical, Universidade de São Paulo, Avenuenida Dr Enéas Carvalho de Aguiar, 470, CEP 05403-000, São Paulo, Brasil.

[H]Blood Safety, Hepatitis, STI & HIV Division, National Infection Service, Public Health England, NW9 5EQ, UK.

[I]Corresponding author. Email: katy.turner@bristol.ac.uk

[*]Joint first authors.

The following supporting material is composed of two parts. In section S1, a technical description of the individual based model is provided, following the guidelines of (O)verview, (D)esign concepts and (D)etails established by Grimm et al. (1). Section S2 describes the procedures carried out to calibrate and parametrise the model, in particular how to update the partnership network, and how to set the transmission rate β. The model code is available online: https://data.bris.ac.uk/data/dataset/3erdo698eboli2ptxi324rsuhg

S1 Technical description of individual-based model

Overview

Purpose

The model is a stochastic, discrete-time Markov model which describes individual-level gonorrhoea transmission and recovery, within a dynamic sexual partnership network. Two independent infection strains are modelled, differing only in their response to a choice of two treatment interventions: a ciprofloxacin and ceftriaxone susceptible (non-AMR) strain, and a ciprofloxacin resistant but ceftriaxone susceptible (AMR) strain. Recovery from either infection is either spontaneous (natural recovery), or by means of an intervention in which a choice of two drugs is administered: ciprofloxacin (*Cipr*) or ceftriaxone (*Ceft*). While this is clearly a significant simplification of disease epidemiology and treatment in reality, the aim of the model is to explore whether antibiotic stewardship can be improved by re-introducing an older drug to treat infections to which they are susceptible. The choice of drug and the specific recovery pathway of an individual depends on a number of different experimental scenarios — described in *Implementation of treatment scenarios*. Recovery

pathways are voluntary treatment seeking for symptomatics, partner notification (tracing) and patient recall following misdiagnosis or positivity after random testing. The model framework subsequently allows for different diagnostic and treatment scenarios to be simulated such that we can assess the impact of rapid (point-of-care) and other testing strategies on infection prevalence, and the volume and efficacy of each drug choice.

State variables and scales

Individuals are represented by a set of time varying state vectors representing the infection status with respect to the two gonorrhoea strains, the presence of symptoms, and 'notification' flags controlling if and when individuals attend for diagnostics or treatment. These notifications are used to control the temporal flow of individuals along the discrete treatment and recovery pathways prescribed in the model. They account for differential treatment of symptomatic and asymptomatic infections; the symptom onset period; laboratory test turnaround time; appointment delays for individuals to be recalled for treatment after testing; and similar delays for traced individuals (notified partners) to attend for diagnosis or treatment.

Closed populations of individuals are considered, where the contact network structure and model parameters reflect male men who have sex with men (MSM) interactions. The age and gender of individuals are therefore not considered within the current framework.

Parameter description	**Symbol**	**Value**	**Source**
Partnership network			
Maximum # partners (full network)	k_{max}	120	(2)

Power-law slope	α	1.6	(2)
Restricted degree (max. per update)	k_r	10	Selected
Network update period	τ_r	7 days	Selected
Transmission dynamics			
Transmission rate	β	$2.2\times10^{-3}\ day^{-1}$	Fitted
Natural recovery rate	R	$6.8\times10^{-3}\ day^{-1}$	Fitted
Birth/death rate	μ	$4.6\times 10^{-5}\ day^{-1}$	16-75 years old
Testing rate	γ	$2.5\times 10^{-3}\ day^{-1}$	Fitted
Tracing efficiency	ψ	0.1	Fitted
Symptomatic proportion	P_{symp}	0.5	Fitted
Symptomatic treatment seeking proportion	P_{seek}	0.66	Fitted
Proportion randomly treated with *Ceft*	P_{ceft}	1.0	Guidelines
Behavioural, treatment, and testing delays			
Symptom onset delay	$\Delta_V = \delta_V \pm \epsilon_V$	5 ± 1	Fitted
Voluntary treatment seeking delay	$\Delta_S = \delta_S \pm \epsilon_S$	10 ± 2	Fitted
Recall attendance delay	$\Delta_R = \delta_R \pm \epsilon_R$	2 ± 0.5	Fitted
Trace attendance delay	$\Delta_T = \delta_T \pm \epsilon_T$	7 ± 1	Fitted
(Lab) strain phenotype test result delay	$\Delta_L = \delta_L \pm \epsilon_L$	10 ± 1	Fitted

Table S1: Parameter values and descriptions. Behavioural, treatment, and testing delay parameters are given as mean ± standard deviation of the underlying distribution; see Section S1 for details.

A summary of the state variables for each individual is provided below, in which explicit time-dependence is omitted; the updating protocol is described in section *Process overview and scheduling*. Baseline parameter values for those described can be found in Table S1.

- **Current infection status** Ω: a two-vector for each individual, i, with binary elements indicating the presence or absence of a current infection with strain s, where $s = 1$ is, the non-AMR strain susceptible to both *Cipr* and *Ceft*, and $s = 2$ is the AMR strain susceptible only to *Ceft* treatment. That is,

$$\Omega_{i,s} = \{0,1\}, \qquad s \in \{1,2\},$$

where $\Omega_i = [0\ 0]$ means susceptible to both strains, $\Omega_i = [1\ 0]$ corresponds to infected by the non-AMR strain and susceptible to the AMR strain, $\Omega_i = [0\ 1]$ to infected by the AMR strain and susceptible to the non-AMR strain, and $\Omega_i = [1\ 1]$ to co-infection.

- **Symptoms** $\boldsymbol{S}$: a single binary flag indicating whether an individual i, infected with either strain develops symptoms ($S_i = 1$) or is asymptomatic / not infected ($S_i = 0$). When infected with either strain, individuals are randomly assigned symptoms (which emerge after a delay as described below) with probability P_{symp} or no symptoms with probability $1 - P_{symp}$.

- **Voluntary treatment seeking** $\boldsymbol{seek}$: a single value indicating the simulation day on which a selected symptomatic individual i will seek clinical treatment, having acquired an infection on day T, such that

$$seek_i = T + N_0(\delta_V, \epsilon_V) + N_0(\delta_S, \epsilon_S),$$

where $N_0(\delta, \epsilon)$ is a random normal variate truncated at zero, parametrised by the mean δ and standard deviation (s.d.) ϵ. The parameters δ_V, ϵ_V (respectively, δ_S, ϵ_S) correspond to the mean and standard deviation of treatment seeking delay (respectively, symptom onset delay). Symptomatic individuals seek treatment with probability P_{seek}, decided at the time of new infection. Individuals who are uninfected, asymptomatic or chose not to voluntarily seek treatment for the current infection have $seek_i = \emptyset$.

- **Recall notification $\boldsymbol{recall}$**: a three-vector containing elements (1) the day on which a notification for an individual to return for treatment was issued, 'post-dated' to include the laboratory test turnaround delay $\left(recall_{i,1}\right)$, (2) the day on which the individual will attend for treatment $\left(recall_{i,2}\right)$, and (3) a flag indicating the AMR status of the infection on the day of issue $\left(recall_{i,3}\right)$. The vector as a whole is given by

$$recall_i = \left[T + N_0(\delta_L, \epsilon_L), T + N_0(\delta_L, \epsilon_L) + N_1(\delta_R, \epsilon_R), \Omega_{i,2}(T)\right]$$

where T is the simulation day on which testing / testing was performed, with $N_0(\delta, \epsilon)$ and $N_1(\delta, \epsilon)$ given by random normal variates truncated at zero and one respectively, parametrised by the laboratory test turnaround delay (mean δ_l, s.d. ϵ_l), and patient recall delay (mean δ_r, s.d. ϵ_r). For individuals with no active recall notifications $recall_i = \emptyset$.

- **Trace notification $\boldsymbol{trace}$**: a three-vector containing elements (1) the day on which a partner notification request was 'sent' to individual i- post-dated to include the lab turnaround delay ($trace_{i,1}$), (2) the day on which the notified partner will attend for diagnosis/treatment ($trace_{i,2}$), and (3) the AMR status of the *index* individual j which initiated the notification ($trace_{i,3}$). The vector is given by

$$trace_i = \left[T + N_0(\delta_L, \epsilon_L), T + N_0(\delta_L, \epsilon_L) + N_1(\delta_T, \epsilon_T), \Omega_{j,2}(T)\right]$$

parametrised similarly to $recall$ above but with unique values δ_T and ϵ_T, respectively the mean and s.d. of the delay for notified ($trace$) partners to attend for diagnosis/treatment. For individuals with no active trace notifications $trace_i = \emptyset$.

Transmission dynamics between individuals, and partner notification, are governed by an explicit, time-varying partnership network which is described by the adjacency matrix $A(t) = \{a_{i,j}\}^N{}_{i,j=1} \in \mathbb{Z}^{N\times N}$ – with elements indicating the current presence $(a_{i,j} = 1)$, or absence $(a_{i,j} = 0)$ of a partnership between individuals i and j (see section *Networks modelling of the sexual partnerships*). Susceptible individuals who are not infected with a particular strain at time (day) t can acquire one or both strains simultaneously from an infected partner according a stochastic process and the transmission rate β. Strains transmit independently of each other; there is no change in susceptibility given current (or historical) infection state.

A set of *local* counter variables are employed to track the number of infections, and number and type of drug doses administered to each individual each day. Additional *global* counters (accumulators) are used to quantify the total numbers of each drug (*Cipr* or *Ceft*) administered, further categorised by the infection state of the individual at the time of treatment. These counters therefore provide important metrics describing the suitability of each treatment given the infection state. Other counters include totals for the number of individuals who are screened, traced (partners notified per index case), and recalled at each time step – sub-categorised by the infection strain. The strain-specific incidence (number of new infections per day), and prevalence (number of infected individuals), are also computed after each daily update.

The time horizon for analysis is 1 year, however the model also permits simulations over longer periods, in order to compare how strain prevalence may diverge over longer time periods (see, for example, Figure S1).

Process overview and scheduling

State variables (infection states, notifications, flags and counters) are updated in discrete time, once per day, via a series of distinct sub-processes which iteratively compute the individual state vectors $\Omega_{i,s}(t), S_i, seek_i(t), recall_i(t)$ and $trace_i(t)$ from the corresponding values on the previous day $(t-1)$. All variable updates are performed simultaneously on state vectors, expressed as matrix of length N (e.g. binary, two-strain infection state given by the $\Omega(t)\ \epsilon\ {\mathbb{Z}_2}^{N\times 2}$ 2-vector). Local and global level counters are updated accordingly, either within a sub-process or after all sub-processes are completed. Depending on memory restrictions, state values can either be stored for every simulation day, or simply overwritten at the end of each update step

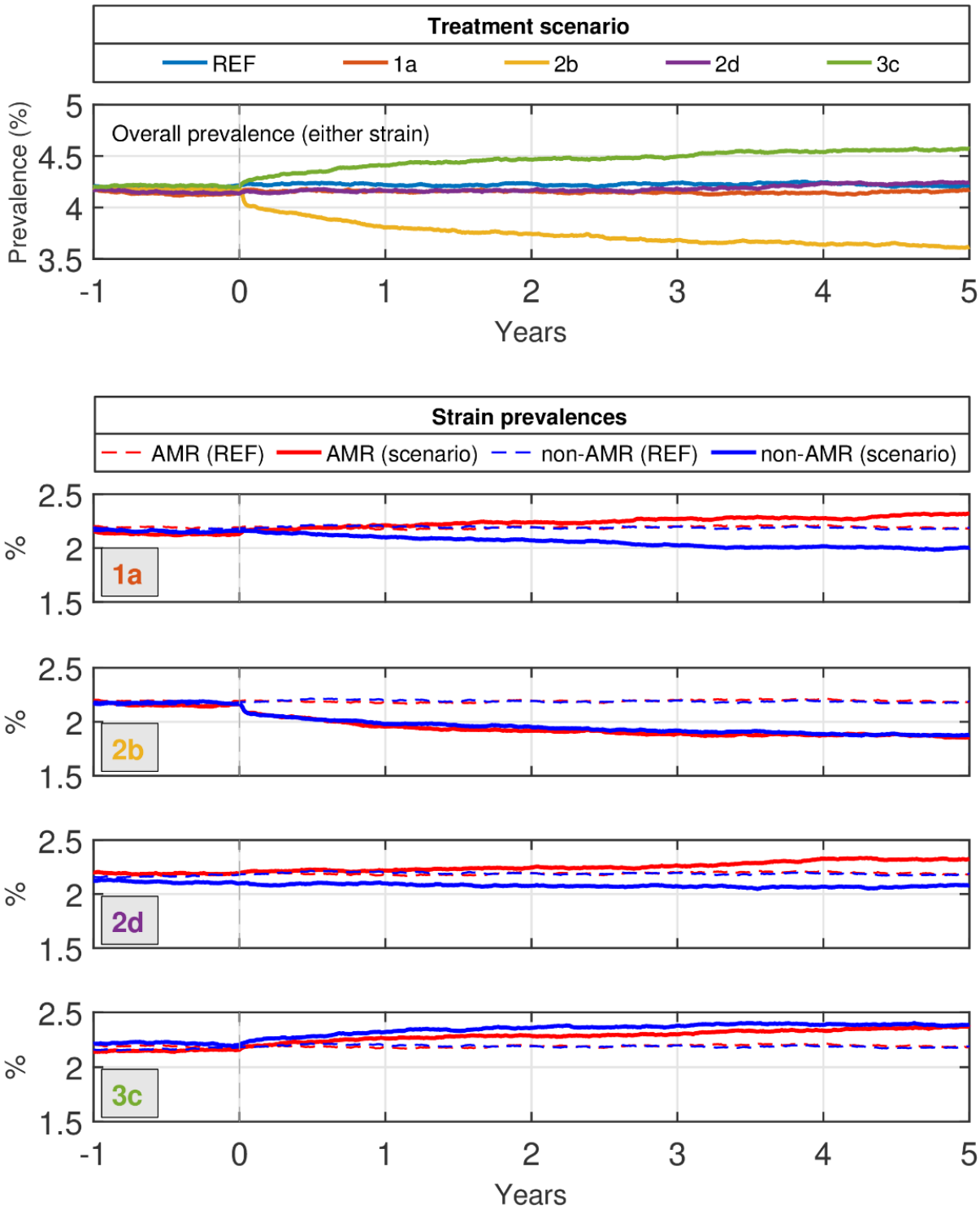

Figure S1 Time-series plots of strain prevalence for different scenarios, with (REF) Reference scenario with 100% Ceft, (1a) 86% Ceft:14% Cipr undirected treatment, (2b) strain discriminatory POCT and targeted treatment, (2d) Informed treatment choice for symptomatic treatment seeking, recalled and traced individuals, (3c) Pre-treatment testing and targeted treatment for traced partners, and 100% Ceft for other, non-screened treatment seekers

A chronological summary of the sub-process executed on each day is given below (further technical details are provided in *Submodels* where indicated):

i. *Partnership network restriction*: (see *Submodels*) The parameter τ_r specifies the period, in days, after which a new restricted, or transient contact network is generated from the initial template partnership network generated on initialisation of each simulation. If τ_r days have passed since the last update, then a restriction algorithm (method described in section *Networks modelling of the sexual partnerships*), is used to generate a new partnership network $A(t)$ with an upper limit of k_r partners per individual.

ii. *Transmission dynamics*: (see *Submodels*) Using the existing infection state vector $\Omega(t-1)$, new infections of either strain are generated using a stochastic process, based on the susceptibility of individuals, and the number of infected partners prescribed by the connectivity matrix $A(t)$. A random proportion P_{symp} of newly infected individuals are assigned symptoms ($S_i = 1$), of which a proportion P_{seek} are voluntary treatment seekers and assigned a date in the future on which they will 'attend' for treatment. Counters and accumulators for infection incidence are updated accordingly.

iii. *Testing*: A random proportion of (all) individuals, with independent probability γ (rate per day), are selected for testing. Individuals currently infected with either strain are given a recall notification including the day of testing, a date in the future on which they will attend for treatment (chosen stochastically, according to the rules described above), and a flag indicating the presence of any AMR infection.

iv. *Natural recovery*: A random proportion of infections present at the previous time step are spontaneously recovered with rate r (per day). This is performed independently for each strain such that a co-infected individual is likely to remain infected with the single remaining strain. Any existing seek, recall or trace notifications remain unchanged.

v. *Births/deaths*: A random proportion of individuals are spontaneously recovered of all infections, and existing seek, recall and trace notifications cleared, with rate μ (per day), remaining in the same network position. This aims to capture a (slow) recycling of the population due to sexual partnership network arrivals and departures. A more literal implementation of this process would add and remove nodes, and associated edges, in the network, while maintaining the overall power-law slope and maximum contact number. Here we consider the size of the network to be fixed over the timescale of the simulation, and incorporate the change in contact structure into the network restriction algorithm that models shorter-time variation (over days or weeks) in the partnership network (see *Submodels*). Since the rate μ is typically very small, we expect the impact on the results of this modelling simplification to be minor.

vi. *Treatment seeking and partner notification*: (see *Submodels*) Individuals who have a notification 'activated' on day t, that is: $seek_i = t$, or $recall_i = t$, or $trace_{i,2} = t$, are selected for treatment according to the current diagnosis and treatment scenario (see *Simulated treatment scenarios*). For example, symptomatic individuals, a proportion of whom choose to seek treatment after symptom onset and seek delays, will enter this treatment pathway on day t if $seek_i = t$, and similarly for those who have been screened, recalled for treatment after previous mistreatment, or notified (traced) by a

partner and flagged to attend treatment on day t. Note that these individuals are not necessarily infected on day t; they may have recovered since being screened or traced, or may not have become infected by contact with an infected partner.

The procedures which follow attending treatment are subject to the specifics of the selected treatment scenario. In general, individual attendees are further selected for treatment with one of the two drug choices – either chosen in an undirected manner, or targeted for treatment according to the AMR status flags $recall_{i,3}$ or $trace_{i,3}$ if present and initialised prior to the current day. Depending on the scenario, individuals who attend via recall and trace pathways may be excluded from treatment and instead subject to (pre)testing.

For those remaining individuals which are treated, a new infection state is generated depending on their current infection state $\Omega_i(t)$ and the efficacy of the drug used, resulting in the following transitions.

- *Cipr* (non-AMR) treatment: reverts any non-AMR infection state $\Omega_{i,1}$ to zero (susceptible), such that

$$\Omega_i(t+1) \longleftarrow \left[0 \;\; \Omega_{i,2}(t)\right]$$

  where an existing AMR strain infection remains unaffected.

- *Ceft* (AMR) treatment: reverts any non-AMR $\Omega_{i,1}$, and AMR $\Omega_{i,2}$ infection states to zero (susceptible), such that

$$\Omega_i(t+1) \longleftarrow [0 \;\; 0]$$

  clearing both infections if either existed prior to treatment (full cure).

Note that for the purposes of this investigation, both drugs are assumed to be 100% effective in treating susceptible strains, to be taken by all individuals to whom they are prescribed, and to clear infections immediately upon prescription. The resistance of the AMR strain $\Omega_{i,2}$ to treatment with *Cipr* is the only differentiating factor between otherwise independent and identically transmittable strains.

Patient recall and partner notification are implemented on the basis that all treatment-seeking individuals will be tested for infection, yielding an immediate result for point-of-care tests, or after a (stochastic) laboratory turnaround delay.

vii. *Final update*: State variables Ω are updated / overwritten to reflect new infections and recoveries. Previously symptomatic individuals who have been recovered are reset to be asymptomatic ($S_i = 0$). Global counters for strain specific prevalences and time (day) counter are incremented, and the simulation procedure repeated from step i., until t exceeds the prescribed time horizon.

Design Concepts

Emergence: Overall and strain-specific prevalences and drug dosing dynamics emerge from the behaviour of the individuals in the model, as described in the overview above; transmission and prescribing are represented by the empirical rules described above, and no "top-down" control (such as, for example, clinic capacity, or drug availability) is applied. Adaptation and fitness-seeking are not modelled.

Sensing: Individuals are assumed to know their neighbours in the network (i.e., with whom they have sexual relationships), although they are not assumed to have perfect recall for the partner notification purposes, embodied in the tracing efficiency parameter. Prescribing

decisions are made either in absence of knowledge of infection type, or with full or partial knowledge when based on diagnostic testing (whether point of care or laboratory-based), dependent on the scenario as described in the Details section below.

Interaction: Individuals interact to transmit the disease, with SIS-type dynamics: infected individuals transmit their disease to susceptible individuals (when connected in the time-varying sexual contact network) at a fixed rate, independent of strain and time. Infected individuals recover, and become susceptible again, as a result of treatment or natural recovery, in a manner that does not depend on any interactions.

Stochasticity: The model in inherently stochastic, to account for natural variability in all the processes we consider. Rates are implemented through Bernoulli processes with probability as specified, while delays are sampled from distributions as specified in the Overview and Details sections. Probabilities and distributions remain fixed, and independent of time and state.

Observation: For model testing and calibration we compute measures of network, disease, and prescribing dynamics. Specifically, we compute overall degree distributions and summary statistics, together with prevalence and incidence of the disease, and drug dosages, both as a function of time and over the duration of the simulation. Prevalence is recorded overall and by strain, while prescriptions are counted in total and sub-divided by efficacy (optimal, over-, under-, and wasted treatment). For model analysis, only disease and prescribing over one simulated year are described.

Details

Initialisation

For all simulations presented in this study, infection dynamics are simulated on populations of $\boldsymbol{N = 10,000}$ individuals with model parameters prescribed as per Table S1. At time (day) $\boldsymbol{t = 0}$, individuals are connected with a random power-law partnership network, derived from the full contact network after one iteration of the degree restriction algorithm described in section *Network modelling of the sexual partnerships*, resulting in a maximum of $\boldsymbol{k_r \leq 10}$ initial partners per individual.

Initial infection states $\Omega_{i,s}(t = 0)$ are generated as identically distributed independent Bernoulli random variates to obtain an overall population prevalence $p_0 = 0.15\ (15\%)$, with equal non-AMR strain (s=1) and AMR strain (s=2) prevalences: ${p_0}^{AMR} = {p_0}^{non-AMR} = p_0/2$. As such, at $t = 0$ there are no individuals who are co-infected with both strains. For infected individuals (either strain), symptomatic state flags S_i are set randomly to 1 (symptomatic) with probability P_{symp} and 0 (asymptomatic) with probability $1 - P_{symp}$. Notification state flags – seek, recall and trace – are all initialised at $t = 0$ with empty values in all entries.

Prior to realisation of each experimental scenario, equilibration periods of up to 10,000 days are implemented, in order to achieve stable infection prevalence and incidence, and to allow infections to travel throughout the population network.

Submodels

Network modelling of the sexual partnerships

The structure of the underlying partnership network has a fundamental impact on the ability of infections to propagate between individuals and throughout an interconnected

population (3). In contrast to vector- or air-borne diseases, the transmission dynamics of those spread via sexual contact, including gonorrhoea, are well described by explicit interaction networks, with structural properties estimated from information on sexual partners. In this study, we derived contact network structure based on observations that the distribution of sexual partners has been reported to follows a power law (2).

Our transmission model simulates infection and recovery of individuals, embedded within an explicit representation of their sexual contacts using a power-law network. In the remainder of this section, we will provide a minimal mathematical description of the network (see *Network notation*) and state explicitly the algorithms used to establish a power-law network (see *Generating random power-law networks*) and the restricted network (see *Contact restriction for high-degree individuals).*

Network notation

Using standard notation (see, for example, (4, 5)), we define a graph, or network $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{1,2, \dots N\}$ is the set of N vertices (nodes) representing individuals, and $\mathcal{E} = \{\{i,j\}: i,j \in \mathcal{V}\}$ is a set of unordered edges (links), or sexual partnerships. Active partnerships between individuals, where $\{i,j\}\, \epsilon\, \mathcal{E}$, can also be expressed conveniently with the $N \times N$ adjacency matrix $A = \left[a_{i,j}\right]$, with elements $a_{i,j} = 1$ if there exists an edge between individuals i and j, and zero otherwise. Here we define a further constraint that sexual partnership networks are *undirected*, i.e. a link from i to j implies an equal and opposite connection from j to i. The adjacency matrix A is therefore symmetric, satisfying the condition

$$A = A^t, \;\; a_{i,j} = a_{j,i}.$$

We also ensure that 'self-loops' are omitted, such that $a_{i,i} = 0$ for all i. The number of active sexual partners, or contact number, of each individual i is given by the degree of each node: $k(i) = |\{i: \{i,j\} \in \mathcal{E}\}|$, equal to the number of links connected to node i. Node degree can equivalently be obtained from the column or row sum of A, for example $k(i) = \sum_{j=1}^{N} a_{i,j}$.

Generating random power-law networks

For each simulated realisation we generate networks where the degree sequence k has a power-law distribution $P(k) = k^{-\alpha}$ with arbitrary slope parameter α, using a Molloy-Reed type degree-based construction method ((6), (7)). To obtain the required *integer* degree sequence, values are rounded to the nearest integer, lower bounded at unity (see (8) for more details on the reliability of this approach).

From the *integer degree sequence*, we generate links between initially unconnected nodes using the following algorithm:

1. Determine the number of free 'stubs' $s(i)$ (residual degree) of each node $i = [1, \ldots, N]$, initially equal to the degree sequence $s = k$.
2. Randomly select a node i with free stubs.
3. Randomly select another node j which has free stubs remaining, i.e. $s(j) > 0$.
4. If i and j are not already connected, add an (undirected) link $\{i,j\} \in \mathcal{E}$, and reduce number of free stubs $s(i)$ and $s(j)$ by one.
5. Repeat from (3) until $s(i) = 0$, or until all potential neighbouring stubs are used, i.e. $\sum_{\{i,j\} \notin \mathcal{E}}^{N} s(j) = 0$.
6. Repeat from (2) until each node has been tested and filled accordingly.

In general, this procedure yields networks with final degree distribution $P(k) \approx P(X)$ where fitted power-law slope parameters α for both distributions are found to be in excellent agreement (Figure S2). Note that compared with alternative algorithms for generating scale-free networks, such as preferential attachment methods (9), this approach allows for flexibility in choosing both α and the degree range $[x_0, x_1]$ – specifically the maximum degree k_{max} – precisely in accordance with available experimental data (2, 7, 10). However, since links between nodes are assigned randomly with no correlation between node degrees, resulting networks have low clustering coefficient (11) and low assortativity (12) compared to those produced via preferential attachment (correlated node degrees). Although expected to be a significant feature of sexual networks (10, 13), clustering is hard to infer or parameterise from data since individuals participating in anonymous surveys typically report only the number of different partners over a given period of time (14-16).

Contact restriction for high-degree individuals

An important component of our model is a novel network algorithm to capture transient partnership dynamics, particularly for highly connected, or promiscuous individuals. Previous models of gonorrhoea transmission dynamics (17, 18), and for chlamydia (19, 20), have focussed on developing extensively parametrised, or data-driven models to capture realistic network dynamics as partnerships form and dissolve over time.

The aim of the network restriction algorithm is to incorporate time variation in partnership networks (timescale of days or weeks) while maintaining the longer-term structure (timescale of years) reported in the literature. We achieve this by producing a new partnership network $\mathcal{G}_r(t)$, at time t, which has maximum (restricted) degree k_r, using the static, annual connectivity network $\mathcal{G}$ containing all possible links as a template. As such,

$\mathcal{G}_r(t)$ is a time-varying sub-network of $\mathcal{G}$, with edges $\mathcal{E}_r \subset \mathcal{E}$. Additionally, we require $\mathcal{V}_r = \mathcal{V}$ such that the restricted graph has the same number of nodes (individuals) as the full network, although some may be disconnected and have degree zero. More succinctly, the effect of restriction is to build a random sub-network between the same N individuals with existing links from the full network, but where the degree, or contact number, of each individual does not exceed the given threshold $k_r < k_{max}$.

Algorithms to generate these random sub-networks can take one of two primary forms: a *reductive approach*, in which we remove links from the full network until the implied condition $k(i) \leq k_r \; \forall \; i$ is met, or an *additive approach*, in which links are added to an empty or time-invariant set of links until no more can be formed. For the experiments discussed later on, we require $k_r \ll k_{max}$ for each update, therefore in the interest of computational efficiency we found it prudent to implement the latter approach, proceeding as follows:

1. Using the degree sequence k of the full network $\mathcal{G}$, find those nodes in $\mathcal{V}' = \{i | k(i) > k_r\}$, which have degree exceeding the threshold k_r.

2. Partition the reduced set of links in $\mathcal{R}$ into two disjoint subsets: $\mathcal{R}_f = \{\{i,j\} | i,j \notin \mathcal{V}'\}$, containing fixed (time-invariant) links between nodes which both have degree less than or equal to threshold k_r, and $\mathcal{R}_c = \mathcal{E} \backslash \mathcal{E}_f$, the complement set of remaining links connected to one or more nodes with degree $k > k_r$ which will be subject to random addition or removal.

3. Set temporary variable $k_c = \{k_c(i) = 0 \mid i \, \epsilon \, \mathcal{V}'\}$ indicating current degree (zero) of nodes with degree greater than threshold k_r in the full network.

4. Create empty set $\mathcal{R}_n = \emptyset$ for storing new links

5. Chose edge from $\mathcal{R}_c = \{i, j\}$, without replacement: if the current degree of nodes i and j are below threshold k_r, i.e. if $k_c(i) < k_r \; and \; k_c(j) < k_r$, then edge $\{i, j\}$ will be accepted and added to $\mathcal{R}_n$. Otherwise link is rejected and discarded.

6. Repeat from (5) until all remaining links in $\mathcal{R}_c$ have been checked

7. The final link list is obtained from the union of fixed links and newly accepted links, i.e. $\mathcal{N} = \mathcal{R}_f \cup \mathcal{R}_n$.

Note that for efficiency, steps (1-2) above can be precomputed at $t = 0$, prior to simulation, and are not required to be re-evaluated for each subsequent update.

Network transmission dynamics

Infections of both strains are acquired according to independent stochastic processes, determined at each time step by computing the infection force F, as follows:

$$F(t) = 1 - (1 - \beta)^{A(t) \times \Omega(\mathrm{t})}$$

at time t, where the i^{th} row values of the matrix $A(t) \times \Omega(t)$ yield the number of partners of individual i who are infected with the non-AMR and AMR strains in columns 1 and 2 respectively. The per-day transmission rate is β, equal for both strains. Here, columns of F_i are the normalised probabilities of individual i acquiring strain 1 (non-AMR), and strain 2 (AMR), at the next time step. The stochastic transmission process therefore proceeds by updating non-zero (strain susceptible) elements of $\Omega(\mathrm{t})$ according to

$$\Omega_{i,s}(\mathrm{t}+1) = \begin{cases} 1, & F_{i,s}(t) > \mathcal{U}_{0,1}, \\ 0, & \text{otherwise}, \end{cases} \quad \forall\, \Omega_{i,s}(t) = 0$$

where $\mathcal{U}_{0,1}$ are independent, uniform random variates on [0,1]. Note that this equation generates new infections only, such that elements indicating a previous infection with either strain remain unaffected at the next time step. Individuals who are newly infected with either strain are assigned 'symptoms', setting the associated symptoms flag S to 1 with probability P_{symp}. A proportion P_{seek} of these individuals are further labelled as "*symptomatic treatment seeking*" and are uniquely assigned a random day in the future, selected as described in section *State variables and scales* (see *Voluntary treatment seeking*).

Implementation of treatment scenarios

We study a range of different diagnostic and treatment scenarios, in order to determine their relative impact on treatment outcomes, drug usage and resulting strain transmission dynamics (see sections *Methods and Results* in the manuscript). Key inferences are made with respect to a reference scenario (REF) which specifies that *all* individuals who seek treatment, or attend for treatment after a recall or trace notification, on a given day, are prescribed *Ceft*. As such, the reference scenario is intended to model the current status of gonorrhoea treatment in which present guidelines recommend only *Ceft* as first-line treatment (21). Within this scenario, the AMR status flags for subsequent recall and trace notifications are effectively ignored, since treatments are homogeneous and strain independent. A positive test result for either strain, albeit delayed by δl, is therefore sufficient to issue any relevant notifications. Model parameters for simulations of the reference scenario, and all other scenarios unless otherwise specified, are given in Table S1 – selected to obtain incidence and prevalence levels within the estimated target ranges (see *Transmission rate β calibration*).

Additional scenarios (see section *Scenario analysis* in the manuscript), are split into three categories based on the type of diagnostic treatment strategy employed:

1. **Undirected drug choice:** Scenarios (1a) and (1b) consider assignment of drug treatments regardless of strain phenotype identification which might otherwise inform susceptibility to a particular antibiotic. In these scenarios, symptomatic treatment seeking patients, individuals recalled after testing with positive diagnosis (either strain), and all notified partners, are treated in an undirected manner with either *Ceft* or *Cipr*, with fixed likelihood ratio. These scenarios are employed to assess the impact of indiscriminate treatment which does not conform with current guidelines, as has been previously reported (22). Data obtained from these simulations also provides a useful comparison to subsequent scenarios in which off-guideline treatments, i.e *Cipr*, are used with discretion, informed by available phenotype test results. In scenario (1a), 86% of treatments are randomly assigned to be with *Ceft*, $P_{ceft} = 0.86$, reflecting the state of recent UK prescribing where approximatley 86.5% of new diagnoses were treated according to UK guidelines (22). In this model, the remaining 14% of treatments are with *Cipr* and capture potential treatment failures due to resistance. In reality, different drugs or drug combinations may be used, for example Ceftriaxone with doxycycline, or Cefixime with azithromycin. In scenario (1b) we propose a more exaggerated scenario, in which only half of all treatments are with *Ceft*, where $P_{ceft} = 0.5$. This allows us to explore the impact of mass under-treatment (treatment failures). In this, and all simulated scenarios in which treatment with *Cipr* is possible, any individual i infected with the AMR strain and under-treated with *Cipr* generates a recall notification with an AMR

risk flag ($recall_{i,3} = 1$) and will attend for re-treatment after the appropriate delay ($\Delta_L + \Delta_R$).

2. **Individualised treatment:** Scenarios (2a) and (2b) simulate the impact of an idealised (100% accurate) point-of-care test and subsequent treatment. In (2a) we consider non-discriminatory POCT under which treatment seeking, recalled and traced individuals are treated with *Ceft* (only), if and only if they are infected with either (or both) gonorrhoea strains on the day of attendance. In (2b), a discriminatory POCT is simulated such that *Ceft* and *Cipr* are prescribed according to the infection status of the individual – *Ceft* for individuals with an AMR infection component, and *Cipr* to those who do not. Consequently, in both POCT scenarios, no treatments are given to individuals who are uninfected. Scenarios (2c) and (2d), simulate a novel diagnostic and treatment regime, individualised in a way that is currently realisable, in which laboratory strain phenotype test results (subject to a delay of δ_L days) are used to make an informed treatment choice. The first implementation (2c) applies this strategy only to individuals attending on day t who were recalled for treatment after a positive screen for either strain. The default treatment for these individuals, and any traced individuals, is *Ceft* regardless of their actual infection status. If, however, any of these individuals have a recall notification, valid on day t which has the AMR risk flag set to zero ($recall_{i,3} = 0$), *Cipr* is prescribed. Specifically, this flag will be zero if, on the day of testing, no AMR strain was detected. In scenario (2d), the same applies but with additional informed treatment for traced individuals attending on t. Once again, the default treatment for these individuals is *Ceft*, unless the recall *and* trace AMR risk flags are both zero ($recall_{i,3} = trace_{i,3} = 0$). Here, traced individuals

are treated with *Cipr* providing the associated index case was only infected with the non-AMR strain, regardless of what treatment the index case received.

3. **Pre-treatment testing:** Scenarios (3a)–(3d) simulate variations of previous scenarios, with additional degrees of 'pre-treatment testing' – where subsets of individuals attending for treatment are screened, i.e. tested for infection, prior to receiving treatment. The pathway by which these pre-screened individuals receive treatment becomes the same as those attending via standard testing and is dependent on any subsequent recall notifications generated, thus accruing an additional delay due to both the lab test and patient recall delay ($\Delta_L + \Delta_R$). The concept behind these treatment scenarios is to address observed patterns of treatment wastage, for example by assuming traced contacts to be infected, and treating them regardless of whether or not they show symptoms (or, indeed, are infected). Initially in scenario (3a) pre-treatment testing is applied only to secondary cases, i.e. traced partners of index cases previously found to be infected. In scenario (3b), pre-treatment testing is extended to include index cases as well as traced, secondary cases. We note that in reality it would be highly unusual for clinicians to await extensive test results before treating a patient with obvious symptoms, however this scenario provides an interesting comparison to highlight the source of treatment wastage (treatment of uninfected individuals). As per the reference scenario, final treatments for (3a) and (3b) are exclusively with *Ceft*. In scenarios (3c-3d), informed discriminatory treatment is included as per scenarios (2c) and (2d), with additional pre-treatment testing according to (3a) and (3b) above respectively. In these scenarios we therefore allow treatment with *Ceft* as an alternative to *Ceft* in the depending on

strain discriminatory testing results. Specifically, symptomatic and asymptomatic individuals who are attending after being recalled due to a positive diagnosis (either strain) will be treated according to the strain specificity result of the laboratory test – i.e. *Cipr* when no AMR strain detected. However, in (3c) symptomatic individuals who seek treatment prior to any testing results being available will be treated according to current guidelines (*Ceft*). In (3d), all symptomatic treatment seekers are pre-screened and therefore can receive strain specific treatment if and when they are recalled. In this latter scenario therefore, all individuals are treated according to the strain phenotype detected at the time of testing. However unlike with POCT scenarios, there are additional laboratory delays prior to treatment during which individuals may continue to spread infections within the population.

S2 Model calibration

Cumulative partnership networks: mean, median, mode and maximum of degree sequence over one year

Contact network

In order to validate the dynamical partnership process (see *Contact restriction for high-degree individuals*), which periodically shuffles active partnership of high degree individuals, we check whether the cumulative structure of the dynamic network converges sufficiently on the full partnership 'annual template' network. For this analysis, we generated a single template network with $N = 10{,}000$ nodes, a power-law slope parameter $\alpha' = 1.6$, and

maximum degree $k_{max} = 120$ (see Table S1). Time-integration was achieved by computing the element-wise *OR* operator across the binary adjacency matrices generated with each update of the network, over 1 year (365 days) of simulation. Network integration was performed for different values of k_r, the restricted maximum degree, and the network update period τ_r with $\lfloor 365/\tau_r \rfloor$ updates per year. The resulting integrated networks are compared with the original network (with $k_r = k_{max} = 120$), quantified by their degree distribution and power-law slope parameter α. Best fit values for α are calculated via a maximum likelihood procedure(8).

In Figure S2, we report the results of tests for restricted networks with maximum degree $k_r \epsilon \{2,5,10,20,60,120\}$, with update periods $\tau_r \epsilon \{1,7,14,30\}$ days, shown in panels **A-D**. In each example we include the static, unrestricted network with $k_r = k_{max} = 120$ and power law slope $\alpha' = 1.6$ (red) to which we compare the other distributions.

In panel **A**, when $\tau_r = 1$ (daily update), we find the best-fit slope parameter $\alpha = 1.60 \pm 0.005$ to be indistinguishable from the template value α' for all values of $k_r \geq 2$. Furthermore, each of the k_r degree distributions are found to overlap right up to tail where the sharp decrease in gradient indicates finite size effects: truncation at the maximum degree $k_{max} \approx 120$. Even when k_r is as low as 2, daily updates are sufficient to generate an annual integrated network which attains the specified maximum yearly contact number, with the same scale-free structure parametrised by α. However, such rapid updates will lead to information being lost between recent partners, particularly affecting contact tracing. Simulation speed is also found to be heavily impacted by such frequent re-computation of transient network structures.

For weekly updates, $\tau_r = 7$, and highly restricted networks with $k_r = 2 - 5$ in Panel **B**, we find that partner accumulation over one year is insufficient, for individuals with high degree in the full network, to sample all possible partners. This trend continues with fortnightly (Panel **C**) and monthly (Panel **D**) updates, where we find degree distributions for smaller values of k_r to be increasingly truncated with respect to the full network distribution. For example, for monthly updates in panel **D**, highly restricted $k_r = 2$ networks integrate to a maximum degree of only 24, and thus exclude many possible partnerships over the course of a year.

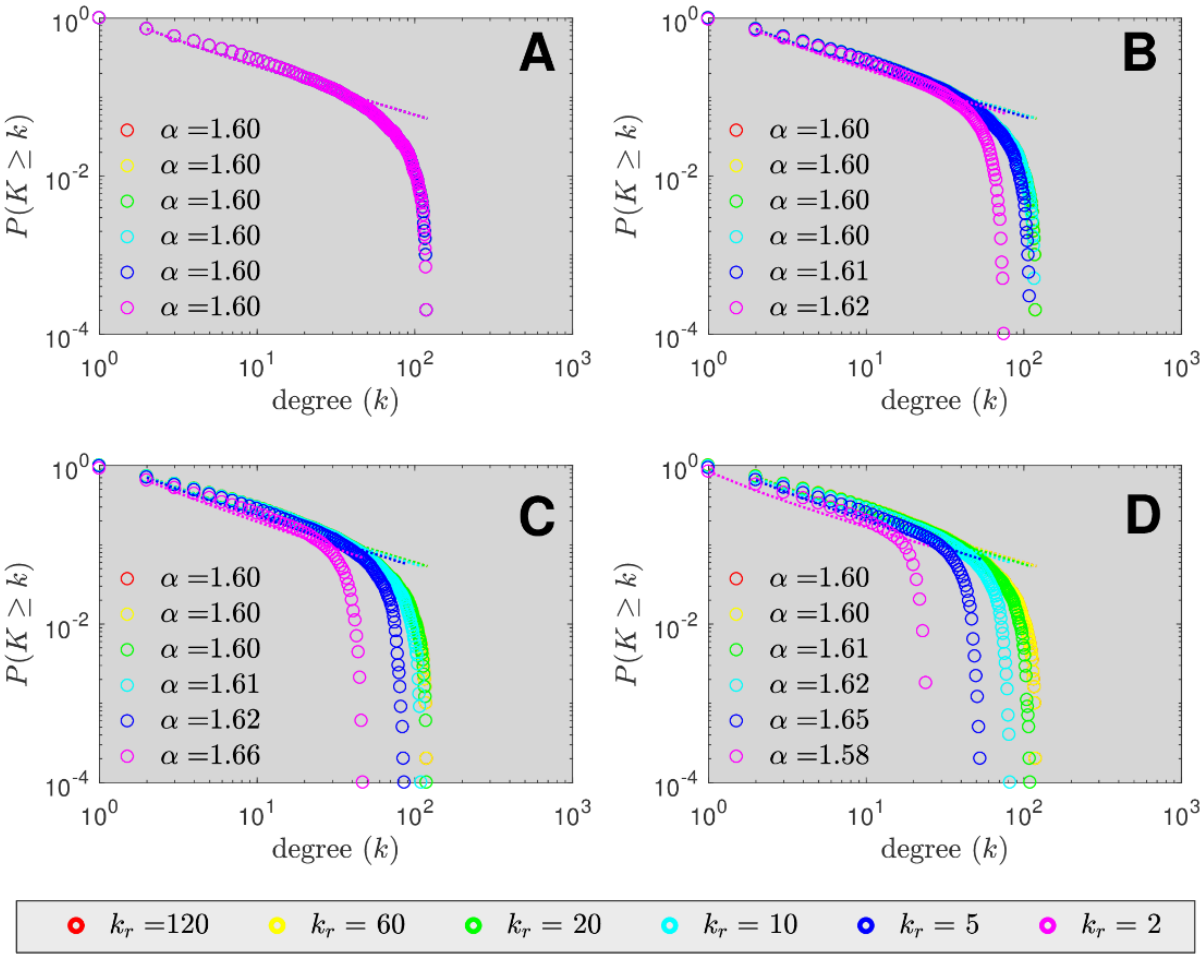

*Figure S2: Degree distribution of restricted sub-networks integrated over 1 year; panels represent different update periods, τ_r (days): **(A)** τ_r=1, **(B)** $\tau_r = 7$, **(C)** $\tau_r = 14$, and **(D)** $\tau_r = 30$. Colours indicate different values of k_r, the maximum degree*

of the restricted sub-network. Plotted on log-log axis, the power-law distributions of the degree sequences are characterised by a straight line with slope α (dotted line), fitted value as indicated in the legend.

Across all combinations of k_r and update period τ_r, we find only minor variation to the fitted slope parameter α. Largest deviations (up to $\alpha = \alpha' + 0.06$) are found for highly restricted (low k_r) networks, exacerbated when the update rate is low (high τ_r). Integrated networks of low maximum degree, derived from the full template, therefore display similar scaling behaviour; however, the degree distribution and range become unacceptably truncated with respect to the desired final state after one year.

For each parameter combination, we also computed the mean, median, mode and range of the resulting degree sequences. The results are shown in Figure S2 (with red dotted line showing the expected result for the full partnership network). From these extra indicators, we find that the cumulative degree truncation, due to low k_r and increased τ_r, is found to be largely mitigated for values of $k_r \geq 10$. The parameters chosen for this study ($\tau_r = 7, k_r = 10$) yield excellent agreement to the average (mean/median/mode) contact number expected after one year, and reasonable agreement with the maximum contact number, while preserving power law slope, and reducing computational effort.

Figure S3 compares the contact number distribution, resulting from integrated networks with $\tau_r = 7, k_r = 10$, to available UK survey data for MSM (14). In general, we find satisfactory agreement across the ranges specified in the survey apart with the exception of surveyed individuals reporting no sexual partners in the previous year. Random, integrated networks in our model however yield a very low probability of un-partnered individuals, and correspondingly overestimate the proportion who have a single partner.

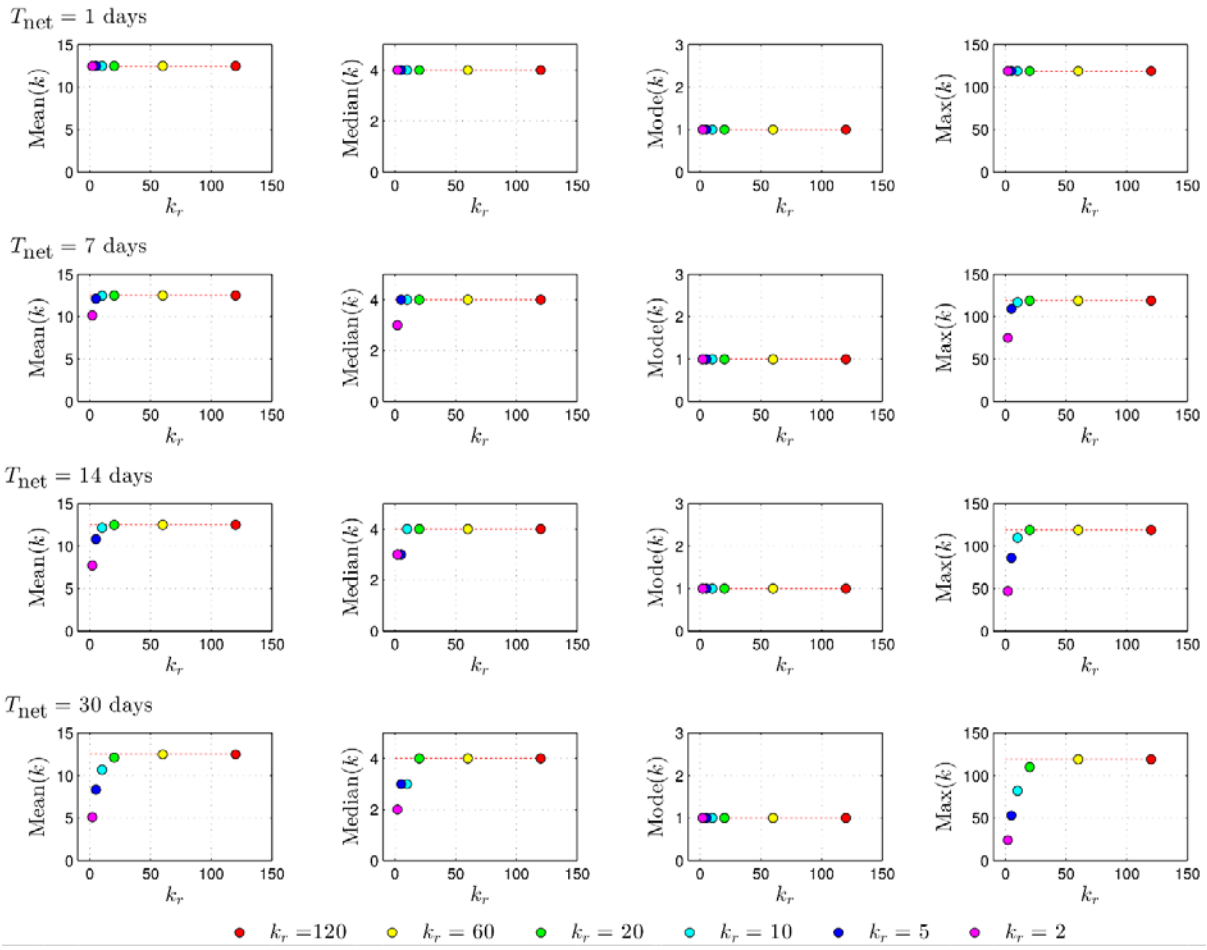

Figure S3: Mean, median, mode and max (columns) of the annual integrated degree distribution, as a function of restricted maximum degree k_r (values as in the legend at the bottom of the figure), for different updating periods $\tau_r = 1, 7, 14, 30$ days (rows).

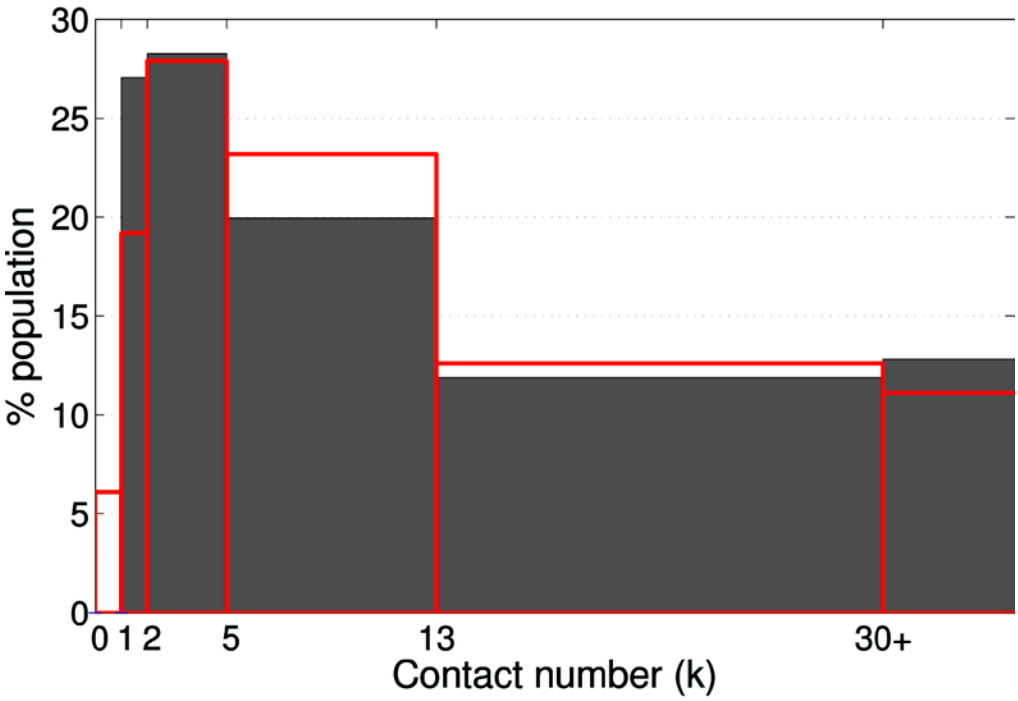

Figure S4: Annual contact number over one year as a proportion of MSM population: comparing integrated model contact network with update period $\tau_r = 7$ and maximum degree $k_r = 10$ per update (gray bars), with data from (14)

Transmission rate β calibration

In what follows we report sensitivity data for selected model parameters with variable transmission rate, β, sampling end-of-simulation values for infection prevalence of a single (non-AMR) infection strain, and the number of positive diagnosis per day (per 10,000 individuals). The objectives of this analysis are to determine the relative sensitivity of key parameters and, more specifically, to determine a suitable transmission rate which yields equilibrium prevalence and diagnosis rates within realistic ranges for the target population of MSM in London, U.K.

Panels in Figure S5 consist of contoured heat-maps in which colour indicates the equilibrium prevalence (left column), and positive daily diagnosis rates (right column) – averaged over the final 30 days of simulation for each parameter combination. Each row presents this

output data for different model parameters on the vertical axis, and transmission rate β increasing on the horizontal axis. Regions of parameter space which yield target prevalence (3-10%), and daily positive diagnoses (2-5 per 10,000 individuals) are shaded light-blue and red respectively. Other parameters, held constant, are as given in Table S1. In order to achieve the target prevalence and the number of positive diagnoses simultaneously, we overlap the shaded regions in each case (left and right columns in Figure S5) and consider the value of $\beta = 0.0022\ day^{-1}$ at the intersection as our fitted value for β in Table S1.

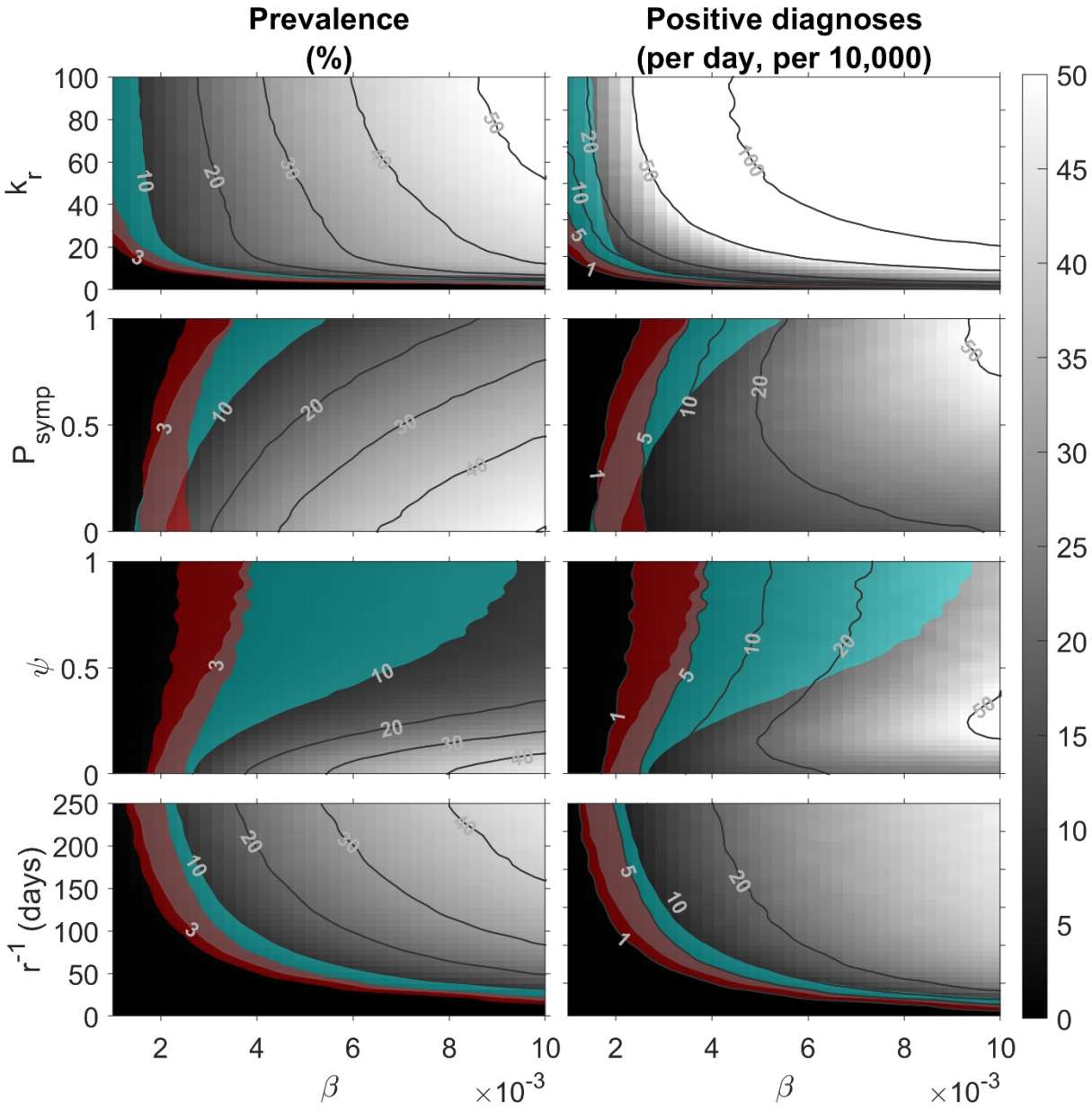

Figure S5: Bivariate parameter sensitivity of single strain gonorrhoea transmission model: prevalence (left panels), and confirmed positive diagnoses (right panels). From top to bottom we vary: the maximum degree of the restricted sub-networks k_r, the proportion of symptomatic infections P_{symp}, tracing efficiency ψ and the mean untreated infectious duration $1/r$, each as a function of the transmission rate β. Other parameter values are as per the reference scenario, given in Table S1. Superimposed shaded regions indicate parameter space which yields stable target prevalences of between 3-10% (light blue), and daily positive diagnosis rates between 2-5 per 10,000 individuals (red).

References

1. Grimm V, Berger U, Bastiansen F, Eliassen S, Ginot V, Giske J, et al. A standard protocol for describing individual-based and agent-based models. Ecological Modelling. 2006;198(1-2):115-26.
2. Schneeberger A, Mercer CH, Gregson SA, Ferguson NM, Nyamukapa CA, Anderson RM, et al. Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. Sex Transm Dis. 2004;31(6):380-7.
3. Pastor-Satorras R, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. . Epidemic Processes in Complex Networks. Reviews of Modern Physics. 2015;87:925-79.
4. Newman MEJ. The Structure and Function of Complex Networks. SIAM Review. 2003;45(2):167-256.
5. Newman M, Barabasi DJ, Watts A. The Structure and Dynamics of Networks: Princeton University Press; 2006.
6. Molloy MR, B. A Critical Point for Random Graphs with a given Degree Sequence Random Structures & Algorithms. 1995;6 161–80.
.
7. Kim H, Toroczkai Z, Erdos P, Miklos I, Szekely LA. Degree-based graph construction. Journal of Physics A: Mathematical and Theoretical. 2009;42(39):12.
8. Clauset A, Shalizi CR, Newman ME. Power-law distributions in empirical data. SIAM Review. 2009;51(4):661-703.
9. Barabasi AL, Albert R. Emergence of scaling in random networks. Science. 1999;286(5439):509-12.
10. Liljeros F, Edling CR, Amaral LA, Stanley HE, Aberg Y. The web of human sexual contacts. Nature. 2001;411(6840):907-8.
11. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. Nature. 1998;393(6684):440-2.
12. Newman ME. Mixing patterns in networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2003;67(2 Pt 2):026126.
13. Wasserman S, Faust K. Social network analysis: Cambridge University Press; 1994.
14. Hickson F, Weatherburn P, Reid D, Jessup K, Hammond G. Consuming passions: findings from the United Kingdom Gay Men's Sex Survey 2005 2005 [Available from: http://sigmaresearch.org.uk/reports/item/report2007c.
15. Mercer CH, Tanton C, Prah P, Erens B, Sonnenberg P, Clifton S, et al. Changes in sexual attitudes and lifestyles in Britain through the life course and over time: findings from the National Surveys of Sexual Attitudes and Lifestyles (Natsal). Lancet. 2013;382(9907):1781-94.
16. European Centre for Disease Prevention and Control. The European Men-Who-Have-Sex-With-Men Internet Survey. Findings from 38 countries. 2013 [Available from: https://www.ecdc.europa.eu/en/home.
17. Kretzschmar M, van Duynhoven YT, Severijnen AJ. Modeling prevention strategies for gonorrhea and Chlamydia using stochastic network simulations. Am J Epidemiol. 1996;144(3):306-17.
18. Ghani AC, Swinton J, Garnett GP. The role of sexual partnership networks in the epidemiology of gonorrhea. Sex Transm Dis. 1997;24(1):45-56.

19. Turner KM, Adams EJ, Gay N, Ghani AC, Mercer C, Edmunds WJ. Developing a realistic sexual network model of chlamydia transmission in Britain. Theor Biol Med Model. 2006;3:3.
20. Low N, McCarthy A, Macleod J, Salisbury C, Campbell R, Roberts TE, et al. Epidemiological, social, diagnostic and economic evaluation of population screening for genital chlamydial infection. Health Technol Assess. 2007;11(8):iii-iv, ix-xii, 1-165.
21. Bignell C, Fitzgerald M, Guideline Development G, British Association for Sexual H, Hiv UK. UK national guideline for the management of gonorrhoea in adults, 2011. Int J STD AIDS. 2011;22(10):541-7.
22. Public Health England. GRASP 2013 Report: Gonococcal Resistance to Antimicrobials Surveillance Programme 2013 [Available from: http://webarchive.nationalarchives.gov.uk/20140714084352/http://www.hpa.org.uk/webc/HPAwebFile/HPAweb{_}C/1317136030908.

.