

Supporting Information: Efficient modelling of infectious diseases in wildlife: a case study of bovine tuberculosis in wild badgers

Evandro Konzen, Richard J. Delahay, Dave J. Hodgson, Robbie A. McDonald,
Ellen Brooks Pollock, Simon E. F. Spencer and Trevelyan J. McKinley

Contents

A Dataset and sampling period for hidden states	2
B Model	3
B1 Transition probabilities for hidden states	5
B2 Probability distribution for the observation process	6
B3 Priors	7
C Algorithm	7
D Parameter estimation	8
D1 HMC updates	8
D1.1 Updating $\alpha, \lambda, \beta, q, \tau, a, b, c$	10
D1.2 Updating θ_j, ρ_j	15
D2 Gibbs sampling updates	15
D2.1 Updating ϕ_j	16
D2.2 Updating η_ℓ	16
D2.3 Updating ν_S, ν_E, ν_I	16
D3 Metropolis-Hastings update for ξ	17
E Individual effective reproduction number	17
E1 Discrete event times	17
E2 Population-level effective reproduction numbers	19
F Relative rates of infection	20
G Additional results	21
G1 Mixing	21
G2 Additional model outputs	22
G3 Social group model fits	25

A Dataset and sampling period for hidden states

We use data from 34 social groups. The original dataset contains $m = 2751$ individuals, but we only use the individuals which, at their first capture, were found in a core social group. We discard 360 individuals (13.1%), out of which: 165 (6.0%) were never captured in any of the core social groups; 63 (2.3%) were initially captured in a non-core social group and moved to a core social group afterwards; 132 (4.8%) were first captured in a social group which was not included in the model.

Since trapping effort was consistent in some social groups at different times, each social group has its own monitoring period. In addition, we assume the monitoring period for an animal ends if it leaves the core area. We conduct inference on the hidden states of the animals only in their corresponding monitoring periods. However, trapping events occurring outside of the monitoring period are still used in the mortality component of the model. Figure A shows the monitoring periods for each core social group.

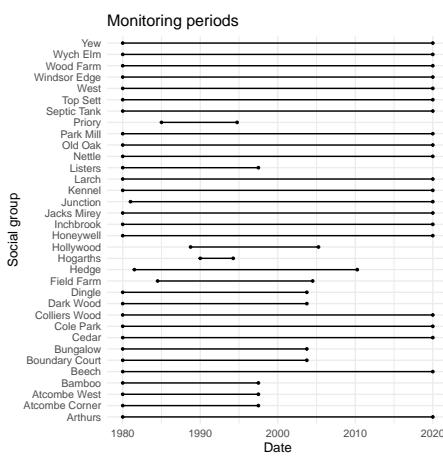


Figure A: Monitoring periods for each core social group.

During each monitoring period, trapping is conducted each quarter. However, trapping effort is determined by first checking all the known setts within a social group to see if they are active (i.e. are there signs of digging, footprints, used entrance holes etc.). If they show signs of activity then traps are deployed, but if they are inactive then traps are not deployed because there is no indication that any badgers are present at that time. The exact numbers of traps deployed are not recorded, but trapping efficiency has been investigated in [11]. Note that social groups usually comprise of more than one sett, which is why we model trapping through a probabilistic process at the social group scale, which incorporates some of these uncertainties. For more information on how trapping is conducted at Woodchester Park, see [e.g. 3, 2, 6].

At the known birth times (recorded at quarterly time intervals), we assume the badgers are not infected. We also assume that individuals belong to a social group until they are captured in a new social group. Badgers were sometimes captured and tested multiple times at the same quarter. All these test results are used and assumed to be independent given the hidden states.

For animals born during the monitoring period, we sample the infection states from their date of birth. For animals born beforehand, we assume they are susceptible, exposed, or infectious with probability ν_S , ν_E , and ν_I , respectively, at the beginning of the monitoring period. Infection states of individual i are sampled up to the end of the monitoring period (T_i).

B Model

Notation

- $\Omega = \{S, E, I, D\}$: set of possible states (susceptible, exposed, infectious, dead)
- N : number of states (number of elements in Ω)
- m : number of individuals
- T : number of time points of the discrete-time hidden state process
- $X_t^{[i]} \in \Omega$: hidden state variable for individual $i \in \{1, \dots, m\}$ at time $t \in \{1, \dots, T\}$
- $\mathbf{X}_t^{[1:m]} = (X_t^{[1]}, \dots, X_t^{[m]})'$: vector of infectious states of m individuals at time t
- $\mathbf{X}_{1:t}^{[1:m]} = (\mathbf{X}_1^{[1:m]}, \dots, \mathbf{X}_t^{[1:m]})'$: set of m individuals at times $1, \dots, t$
- $\mathbf{X}_t^{[-i]}$: vector of all individuals, except i , at time t
- $W_t^{[i]}$: random binary variable for the capture process of individual i at time t
- $\mathbf{Y}_t^{[i]} = (Y_{1t}^{[i]}, \dots, Y_{Jt}^{[i]})$: random vector for the outcomes of the J tests for individual i at time t
- Data:
 - $w_t^{[i]}$: observed binary variable indicating whether individual i is captured at time t : $w_t^{[i]} = 1$ if captured, and 0 otherwise
 - $o_{jt}^{[i]}$: observed binary variable indicating whether test j , $j = 1, \dots, J$, was taken for individual i at time t : $o_{jt}^{[i]} = 1$ if the test was taken, and 0 otherwise
 - $\mathbf{o}_t^{[i]} = (o_{1t}^{[i]}, \dots, o_{Jt}^{[i]})'$: binary vector indicating whether individual i was tested by each of the J tests at time t
 - $y_{jt}^{[i]}$: observed result of test j , $j = 1, \dots, J$, at time t for individual i : $y_{jt}^{[i]} = 1$ if tested positive, and 0 if tested negative
- $\mathcal{I}_{i,t-1}$: set of individuals who could infect individual i in the time period $[t-1, t)$
- $\mathcal{M}_{i,t-1}$: set of individuals in the social group to which individual i belongs in the time period time $[t-1, t)$ (including individual i itself)
- $|\mathcal{X}|$: cardinality of the set \mathcal{X}
- t_i^B, t_i^E, t_i^I : times at which individual i is born, becomes exposed, and becomes infectious, respectively
- $t_{0,i}$: starting time for the inference of hidden states of individual i . It is either the birth time or the start of the monitoring period, whichever is later
- T_i : end of monitoring period for individual i
- T_i^{LC} : last capture time of individual i
- $\psi_t^{[i]}$: probability that individual i dies in $[t-1, t)$ given it is alive at $t-1$.
- $\mathbb{1}_{\{A\}}$ indicator function that equals 1 if A is true, and 0 otherwise
- $I_t^{[i]}(j, l)$: indicator function that equals 1 if individual i is in state j at time $t-1$ and in state l at time t , and 0 otherwise.
- α_g : background rate of infection in the g -th social group, $g = 1, \dots, G$
- $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_G)'$: vector with all the background rates of infection
- $\alpha_{g_{i(t-1)}}$: background rate of infection of the social group that individual i belongs to in the time period time $[t-1, t)$
- λ : parameter used in the hierarchical shrinkage prior on $\boldsymbol{\alpha}$
- a, b, c : Gompertz-Makeham distribution parameters for mortality
- $S(u) := S(u; a, b, c) = P(U \geq u)$: survival function at age u
- β : badger-to-badger transmission rate

- τ : average latent period
- q : parameter to assess degree of density vs. frequency dependence
- ν_S, ν_E, ν_I : initial probabilities of being susceptible, exposed, and infectious, respectively, at the beginning of the monitoring period
- ξ : Brock changepoint
- θ_j : sensitivity of test j during the infectious period
- ρ_j : scaling parameter for the test sensitivity of test j during the latent period, so that $\rho_j \theta_j$ is the sensitivity of test j during the latent period
- ϕ_j : specificity of test j
- η_ℓ : probability of capturing an alive individual at the ℓ -th quarter, $\ell = 1, \dots, 4$.
- s_t : variable indicating which quarter of the year time t corresponds to; formally, $s_t := 1 + ((t - 1) \bmod 4)$, where ‘ \bmod ’ refers to the modulo operation, so that $s_t \in \{1, 2, 3, 4\}$.
- Θ : vector including all model parameters
- K : rescaling constant to make the units of β independent of q
- Probability distribution functions:
 - If $Z \sim \text{Exp}(a)$, $f(z) = ae^{-az}$, $a > 0$, for $z \in [0, \infty)$
 - If $Z \sim \mathcal{N}(a, b)$, $f(z) = \frac{1}{b\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{z-a}{b})^2}$, $a \in (-\infty, \infty)$, $b > 0$, for $z \in (-\infty, \infty)$
 - If $Z \sim \text{Gamma}(a, b)$, $f(z) = \frac{b^a}{\Gamma(a)} z^{a-1} e^{-bz}$, $a, b > 0$, where Γ is the Gamma function, for $z \in (0, \infty)$
 - If $Z \sim \text{Beta}(a, b)$, $f(z) = \frac{z^{a-1}(1-z)^{b-1}}{\text{B}(a, b)}$, $a, b > 0$, where $\text{B}(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$, for $z \in [0, 1]$

Coupled hidden Markov models

We use a coupled hidden Markov model (CHMM) [1] for the disease system. In this model, the unobserved state of individual i at time t (denoted by $X_t^{[i]}$) depends not only on its own previous state, but also on the previous state of all other individuals. In addition, it is assumed that the observed data for individual i at time t only depends on the value of $X_t^{[i]}$.

The joint posterior density for the hidden states and model parameters is given by

$$\begin{aligned} \pi(\Theta, \mathbf{X}_{1:T}^{[1:m]} \mid \mathbf{W}_{1:T}^{[1:m]}, \mathbf{o}_{1:T}^{[1:m]}, \mathbf{Y}_{1:T}^{[1:m]}) \propto \pi(\Theta) \times \\ \left[\prod_{i: t_i^B < t_{0,i}} S(t_{0,i} - t_i^B) P(X_{t_{0,i}} \mid \Theta) \right] \times \\ \left[\prod_{i=1}^m \prod_{t=t_{0,i}+1}^{T_i} P(X_t^{[i]} \mid X_{t-1}^{[i]}, \mathbf{X}_{t-1}^{[-i]}, \Theta) \right] \times \\ \left[\prod_{i=1}^m \prod_{t=t_{0,i}+1}^{T_i} P(\mathbf{Y}_t^{[i]}, \mathbf{W}_t^{[i]} \mid X_t^{[i]}, \mathbf{o}_t^{[i]}, \Theta) \right] \times \\ \left[\prod_{i: T_i^{\text{LC}} > T_i} \prod_{t=T_i+1}^{T_i^{\text{LC}}} (1 - \psi_t^{[i]}) \right], \end{aligned}$$

where the five terms are (i) the joint prior for all parameters; (ii) a survival term times a probability distribution for the states at the beginning of the monitoring period for individuals born beforehand; (iii) the probability distribution for transitions of the hidden states (these include mortality components); (iv) the probability distribution for the observation process; and (v) a survival term for individuals captured after the monitoring period. All these terms are specified below.

B1 Transition probabilities for hidden states

Gompertz-Makeham mortality for $\{S, E, I\} \rightarrow D$ transitions

Let $\psi_t^{[i]}$ be the probability of individual i dying between $t - 1$ and t given it is alive at $t - 1$. This will be

$$\psi_t^{[i]} := \frac{S(\text{age}_t^{[i]} - 1) - S(\text{age}_t^{[i]})}{S(\text{age}_t^{[i]} - 1)} = 1 - \frac{S(\text{age}_t^{[i]})}{S(\text{age}_t^{[i]} - 1)}, \quad (1)$$

where

$$S(u) := S(u; a, b, c) = \exp \left\{ -cu + \frac{a}{b}(1 - e^{bu}) \right\}, \quad u \in [0, \infty), \quad a, b, c > 0, \quad (2)$$

is the continuous-time Gompertz-Makeham survival function.

$S \rightarrow E$ transition probabilities given survival

We first describe the transition probability for a susceptible individual to become infected in the time period $[t - 1, t]$. We assume there are two sources of infection—the background and badger-to-badger transmission. For the latter, we assume that (i) each infectious individual infects a susceptible at rate β ; (ii) all infectious individuals are equally likely to infect susceptibles in the same social group; and (iii) infectious individuals are independent in $[t - 1, t]$. Therefore,

$$F_{g_{i,t-1}} = \alpha_{g_{i(t-1)}} + \beta K^q \frac{|\mathcal{I}_{i,t-1}|}{|\mathcal{M}_{i,t-1}|^q}$$

is the force of infection acting on the susceptible individual i between times $t - 1$ and t , given it survived until time t . The parameter q measures the strength of frequency dependence [10, 4], and K is a rescaling constant that makes the units of β independent of q [10]. The value of K that we used is the median social group size across all social groups and all time.

The probability of no events in $[0, t)$ in a Poisson process with rate Λ is $\exp\{-\Lambda t\}$, motivating the choice of $\exp\{-\Lambda\}$ as the probability of avoiding infection in $[t - 1, t)$ from a source with infection rate Λ . Assuming that all sources of infection are independent, the probability of avoiding infection in $[t - 1, t)$ is

$$\exp\{-F_{g_{i,t-1}}\} = \exp\{-\alpha_{g_{i(t-1)}}\} \left[\prod_{j=1}^{|\mathcal{I}_{k,t_k^E-1}|} \exp\left\{-\beta K^q \frac{1}{|\mathcal{M}_{i,t-1}|^q}\right\} \right],$$

so that the probability that individual i becomes infected in $[t - 1, t)$, given survival until t , is given by

$$P(X_t^{[i]} = E \mid X_t^{[i]} \neq D, X_{t-1}^{[i]} = S, \mathbf{X}_{t-1}^{[-i]} = x_{t-1}^{[-i]}, \Theta) = 1 - \exp\{-F_{g_{i,t-1}}\}.$$

$E \rightarrow I$ transition probabilities given survival

We assume that an exposed individual become infectious at rate $1/\tau$, so that the discrete-time latent period has geometric distribution (given survival) with parameter $p = 1 - \exp\{-1/\tau\}$.

All transition probabilities

Therefore, the transition probabilities for the i -th individual in $[t-1, t)$ are:

$$P\left(X_t^{[i]} = x_t^{[i]} \mid X_{t-1}^{[i]} = x_{t-1}^{[i]}, \mathbf{X}_{t-1}^{[-i]} = x_{t-1}^{[-i]}, \Theta\right) =$$

$$\begin{cases} \left(1 - \psi_t^{[i]}\right) \exp\{-F_{g_{i,t-1}}\} & \text{if } x_{t-1}^{[i]} = S \text{ and } x_t^{[i]} = S \\ \left(1 - \psi_t^{[i]}\right) \left(1 - \exp\{-F_{g_{i,t-1}}\}\right) & \text{if } x_{t-1}^{[i]} = S \text{ and } x_t^{[i]} = E \\ 0 & \text{if } x_{t-1}^{[i]} = S \text{ and } x_t^{[i]} = I \\ \psi_t^{[i]} & \text{if } x_{t-1}^{[i]} = S \text{ and } x_t^{[i]} = D \\ 0 & \text{if } x_{t-1}^{[i]} = E \text{ and } x_t^{[i]} = S \\ \left(1 - \psi_t^{[i]}\right) \exp\{-1/\tau\} & \text{if } x_{t-1}^{[i]} = E \text{ and } x_t^{[i]} = E \\ \left(1 - \psi_t^{[i]}\right) \left(1 - \exp\{-1/\tau\}\right) & \text{if } x_{t-1}^{[i]} = E \text{ and } x_t^{[i]} = I \\ \psi_t^{[i]} & \text{if } x_{t-1}^{[i]} = E \text{ and } x_t^{[i]} = D \\ 0 & \text{if } x_{t-1}^{[i]} = I \text{ and } x_t^{[i]} = S \\ 0 & \text{if } x_{t-1}^{[i]} = I \text{ and } x_t^{[i]} = E \\ 1 - \psi_t^{[i]} & \text{if } x_{t-1}^{[i]} = I \text{ and } x_t^{[i]} = I \\ \psi_t^{[i]} & \text{if } x_{t-1}^{[i]} = I \text{ and } x_t^{[i]} = D \\ 0 & \text{if } x_{t-1}^{[i]} = D \text{ and } x_t^{[i]} \in \{S, E, I\} \\ 1 & \text{if } x_{t-1}^{[i]} = D \text{ and } x_t^{[i]} = D \end{cases}$$

B2 Probability distribution for the observation process

The conditional distribution for the observation process given the hidden states is defined as:

$$P\left(\mathbf{Y}_t^{[i]} = \mathbf{y}_t^{[i]}, W_t^{[i]} = w_t^{[i]} \mid X_t^{[i]} = x_t^{[i]}, \Theta\right) = \begin{cases} P\left(\mathbf{Y}_t^{[i]} = \mathbf{y}_t^{[i]} \mid W_t^{[i]} = 1, X_t^{[i]} = x_t^{[i]}, \Theta\right) & \text{if } w_t^{[i]} = 1, \\ \times P\left(W_t^{[i]} = 1 \mid X_t^{[i]} = x_t^{[i]}, \Theta\right) & \\ P\left(W_t^{[i]} = 0 \mid X_t^{[i]} = x_t^{[i]}, \Theta\right) & \text{otherwise,} \end{cases}$$

since test results cannot exist unless an animal is captured.

The conditional probability distribution of capturing badger i at time t given the hidden states is given by

$$P\left(W_t^{[i]} = w_t^{[i]} \mid X_t^{[i]} = x_t^{[i]}, \Theta\right) = \begin{cases} \eta_{st}^{w_t^{[i]}} (1 - \eta_{st})^{1-w_t^{[i]}}, & \text{if } x_t^{[i]} \in \{S, E, I\} \\ 1 - w_t^{[i]}, & \text{if } x_t^{[i]} = D. \end{cases}$$

The conditional probability distribution for the test results given the hidden states and that the badger is captured is given by

$$P\left(\mathbf{Y}_t^{[i]} = \mathbf{y}_t^{[i]} \mid W_t^{[i]} = 1, X_t^{[i]} = x_t^{[i]}, \Theta\right) =$$

$$\begin{cases} \prod_{j:o_{jt}^{[i]}=1} (1 - \phi_j)^{y_{jt}^{[i]}} \phi_j^{1-y_{jt}^{[i]}}, & \text{if } x_t^{[i]} = S \\ \prod_{j:o_{jt}^{[i]}=1} (\rho\theta_j)^{y_{jt}^{[i]}} (1 - \rho\theta_j)^{1-y_{jt}^{[i]}}, & \text{if } x_t^{[i]} = E \\ \prod_{j:o_{jt}^{[i]}=1} \theta_j^{y_{jt}^{[i]}} (1 - \theta_j)^{1-y_{jt}^{[i]}}, & \text{if } x_t^{[i]} = I \\ 0, & \text{if } x_t^{[i]} = D. \end{cases}$$

Therefore, the resulting joint probability distribution for the observed outcomes given the hidden states will be

$$P\left(\mathbf{Y}_t^{[i]} = \mathbf{y}_t^{[i]}, W_t^{[i]} = w_t^{[i]} \mid X_t^{[i]} = x_t^{[i]}, \mathbf{o}_t^{[i]}, \Theta\right) = \begin{cases} \eta_{s_t} \prod_{j: o_{jt}^{[i]}=1} (1 - \phi_j)^{y_{jt}^{[i]}} \phi_j^{1-y_{jt}^{[i]}}, & \text{if } w_t^{[i]} = 1, \text{ and } x_t^{[i]} = S \\ \eta_{s_t} \prod_{j: o_{jt}^{[i]}=1} (\rho \theta_j)^{y_{jt}^{[i]}} (1 - \rho \theta_j)^{1-y_{jt}^{[i]}}, & \text{if } w_t^{[i]} = 1, \text{ and } x_t^{[i]} = E \\ \eta_{s_t} \prod_{j: o_{jt}^{[i]}=1} \theta_j^{y_{jt}^{[i]}} (1 - \theta_j)^{1-y_{jt}^{[i]}}, & \text{if } w_t^{[i]} = 1, \text{ and } x_t^{[i]} = I \\ 0, & \text{if } w_t^{[i]} = 1, \text{ and } x_t^{[i]} = D \\ 1 - \eta_{s_t}, & \text{if } w_t^{[i]} = 0, \text{ and } x_t^{[i]} \in \{S, E, I\} \\ 1, & \text{if } w_t^{[i]} = 0, \text{ and } x_t^{[i]} = D \end{cases}$$

Let $z_t^{[i]}$ be the Brock test result at time t for individual i in the model without Brock changepoint, and let $y_{1t}^{[i]}$ and $y_{2t}^{[i]}$ be the results for Brock1 and Brock2 tests, respectively, in the model with Brock changepoint. Note that

$$\begin{cases} y_{1t}^{[i]} = z_t^{[i]}, & o_{1t}^{[i]} = 1, \text{ and } o_{2t}^{[i]} = 0, & \text{if } t < \xi \\ y_{2t}^{[i]} = z_t^{[i]}, & o_{1t}^{[i]} = 0, \text{ and } o_{2t}^{[i]} = 1, & \text{if } t \geq \xi \end{cases} \quad (3)$$

B3 Priors

We put a hierarchical shrinkage prior on the background rates of infection:

$$\begin{aligned} \alpha_g &\sim \text{Exp}(1/\lambda), \\ \lambda &\sim \text{Exp}(b_\lambda), \end{aligned} \quad (4)$$

$g = 1, \dots, G$, and set $b_\lambda = 1$. On parameters $x \in \{\beta, \tau, a, b, c\}$, we used a $\text{Gamma}(a_x, b_x)$ prior, where we chose $a_\beta = a_\tau = a_a = a_b = a_c = 1$ (i.e., exponential priors), $b_\beta = b_a = b_b = b_c = 1$ and $b_\tau = 0.01$. On parameters $x \in \{q, \theta_j, \rho_j, \phi_j, \eta_\ell\}$, we used a $\text{Beta}(a_x, b_x)$ prior, where we set $a_x = b_x = 1$ in order to use a uniform prior on all these parameters. On the remaining parameters, we put priors $(\nu_S, \nu_E, \nu_I) \sim \text{Dir}(1, 1, 1)$ and $\xi \sim \mathcal{N}(\mu_\xi = 2000 Q1, \sigma_\xi = 60 \text{ quarters})$.

C Algorithm

Algorithm Pseudo-algorithm of iFFBS for hidden states and MCMC for model parameters.

- 1: Set $r = 1$. Draw Θ from their initial value distributions and initialise $\mathbf{X}_{1:T}^{[1:m]}$ (see the rest of this section);
 - 2: **for** $r = 2, \dots, R$ **do**
 - 3: **for** $i = 1, \dots, m$ **do**
 - 4: Draw $\mathbf{X}_{1:T}^{[i]} \sim P\left(\mathbf{X}_{1:T}^{[i]} \mid \mathbf{X}_{1:T}^{[-i]}, \mathbf{W}_{1:T}^{[i]}, \mathbf{o}_{1:T}^{[i]}, \mathbf{Y}_{1:T}^{[i]}, \Theta\right)$ using iFFBS;
 - 5: **end for**
 - 6: Draw $\Theta \sim \pi\left(\Theta \mid \mathbf{X}_{1:T}^{[1:m]}, \mathbf{W}_{1:T}^{[1:m]}, \mathbf{o}_{1:T}^{[1:m]}, \mathbf{Y}_{1:T}^{[1:m]}\right)$ using appropriate MCMC (see Section D);
 - 7: **end for**
-

For some parameters, we used initial value distributions which were more informative than the corresponding priors: $\lambda, \beta, a, b, c \sim \text{Exp}(100)$, $\tau \sim U(4, 16)$, $\theta_j \sim U(0.5, 1)$, $\rho_j \sim U(0.2, 0.8)$, and $\phi_j \sim U(0.7, 1)$. All the remaining parameters were initially drawn from their priors.

We also had to choose how we initialise the hidden states. For animals born during the monitoring period, we set them to be susceptible on their corresponding date of birth; for animals born beforehand, we set them to be

susceptible, exposed, or infectious, with probability ν_S , ν_E , and ν_I , respectively, at the start of the monitoring period. Each susceptible individual is set to remain susceptible in its lifetime until it is tested positive by any test, at which time we set it to be exposed. The initial latent period for each alive exposed animal is drawn from a geometric distribution with parameter $p = 1 - \exp\{-1/\tau\}$, using the initial value of τ . Finally, the initial death time is set to be the first time at which the individual is no longer captured.

D Parameter estimation

We used Gibbs sampling to update ϕ_j and η_ℓ , random-walk Metropolis Hastings [9] to update ξ , and Hamiltonian Monte Carlo (HMC) [8] to update all the other parameters. Multidimensional adaptive random-walk Metropolis Hastings as discussed in [9] can also be used instead of HMC, but produced slower convergence and mixing.

D1 HMC updates

Notes on HMC for constrained parameters

This section shows how we use HMC to update a constrained parameter x . The idea is to use a symmetrical proposal for the transformed parameter $\tilde{x} = h(x)$ and correct the target distribution.

Suppose the target distribution is

$$\pi_x(x|D) \propto \pi_x(x) f(D|x),$$

where parameter x is constrained. Typical examples are $x \in (0, \infty)$, on which we put Exponential or Gamma prior, and $x \in (0, 1)$, on which we put a Beta prior.

In the HMC, parameters are updated in terms of linear functions of derivatives:

$$x \leftarrow x + \varepsilon \frac{\partial \log \pi_x(x|D)}{\partial x}.$$

Therefore, for constrained parameters, the algorithm can easily go outside the boundaries and we would waste iterations.

We use a transformation of variable approach, so that we propose on the transformed parameter space of $\tilde{x} = h(x)$, which will be on $-\infty < \tilde{x} < \infty$. HMC steps will be of the form

$$\tilde{x} \leftarrow \tilde{x} + \varepsilon \frac{\partial \log \pi_{\tilde{x}}(\tilde{x}|D)}{\partial \tilde{x}},$$

and we need to find $\pi_{\tilde{x}}(\tilde{x}|D)$. Using transformation of variables,

$$f_{\tilde{x}}(\tilde{x}) = f_x(h^{-1}(\tilde{x})) |J|,$$

where $J = \frac{\partial x}{\partial \tilde{x}}$ is the Jacobian of the transformation.

The target distribution (i.e., the posterior of \tilde{x}) is therefore

$$\begin{aligned} \pi_{\tilde{x}}(\tilde{x}|D) &\propto \pi_x(h^{-1}(\tilde{x}) | D) |J| \\ &\propto \pi_x(h^{-1}(\tilde{x})) f(D | h^{-1}(\tilde{x})) |J|. \end{aligned} \tag{5}$$

The rest of this subsection will show the logarithm of the prior and Jacobian terms in Eq. (5), i.e. $\pi_x(h^{-1}(\tilde{x})) |J|$, as well as the derivatives with respect to the transformed parameter \tilde{x} , for different parameter constraints used in the paper. The likelihood terms will appear again when we discuss the posterior distributions.

Prior and Jacobian terms for parameters on $(0, \infty)$

For inferring a positive parameter $x > 0$, we use the logarithmic transformation $\tilde{x} = \log(x)$, and therefore the prior times Jacobian in Eq. (5) is given by

$$\pi_x(h^{-1}(\tilde{x})) |J| = \pi_x(e^{\tilde{x}}) e^{\tilde{x}}.$$

Note that $\pi_x(\cdot)$ is the density function used as the prior for x on the natural scale. The logarithm of the above term is therefore

$$\log \{ \pi_x(h^{-1}(\tilde{x})) |J| \} = \log \{ \pi_x(e^{\tilde{x}}) \} + \tilde{x}.$$

Therefore, by using a Gamma prior $\pi_x(\cdot) = \text{Gamma}(a, b)$,

$$\begin{aligned} \log \pi \{ \pi_x(h^{-1}(\tilde{x})) |J| \} &= \log \left\{ e^{(a-1)\tilde{x}} e^{-be^{\tilde{x}}} \right\} + \tilde{x} \\ &= (a-1)\tilde{x} - be^{\tilde{x}} + \tilde{x} \\ &= a\tilde{x} - be^{\tilde{x}}, \end{aligned}$$

and

$$\frac{\partial \log \{ \pi_x(h^{-1}(\tilde{x})) |J| \}}{\partial \tilde{x}} = a - be^{\tilde{x}}.$$

Prior and Jacobian terms for parameters on $(0, 1)$

For a parameter $x \in (0, 1)$, we use the logistic transformation $\tilde{x} = h(x) = \log \left(\frac{x}{1-x} \right)$, whose inverse transformation is $x = h^{-1}(\tilde{x}) = \frac{e^{\tilde{x}}}{1+e^{\tilde{x}}}$. In this case, the prior times Jacobian term in Eq. (5) is given by

$$\pi_x(h^{-1}(\tilde{x})) |J| = \pi_x \left(\frac{e^{\tilde{x}}}{1+e^{\tilde{x}}} \right) \frac{e^{\tilde{x}}}{(1+e^{\tilde{x}})^2}.$$

The logarithm of the above expression is thus

$$\log \{ \pi_x(h^{-1}(\tilde{x})) |J| \} = \log \left\{ \pi_x \left(\frac{e^{\tilde{x}}}{1+e^{\tilde{x}}} \right) \right\} + \tilde{x} - 2 \log(1+e^{\tilde{x}}).$$

Therefore, by using a Beta prior $\pi_x(\cdot) = \text{Beta}(a, b)$,

$$\begin{aligned} \log \{ \pi_x(h^{-1}(\tilde{x})) |J| \} &= \log \left\{ \left(\frac{e^{\tilde{x}}}{1+e^{\tilde{x}}} \right)^{a-1} \left(1 - \frac{e^{\tilde{x}}}{1+e^{\tilde{x}}} \right)^{b-1} \right\} + \tilde{x} - 2 \log(1+e^{\tilde{x}}) \\ &= (a-1) [\tilde{x} - \log(1+e^{\tilde{x}})] + (b-1) [-\log(1+e^{\tilde{x}})] + \tilde{x} - 2 \log(1+e^{\tilde{x}}) \\ &= a\tilde{x} - (a+b) \log(1+e^{\tilde{x}}) \end{aligned}$$

and

$$\frac{\partial \log \{ \pi_x(h^{-1}(\tilde{x})) |J| \}}{\partial \tilde{x}} = a - (a+b) \left(\frac{e^{\tilde{x}}}{1+e^{\tilde{x}}} \right).$$

Prior and Jacobian terms for the hierarchical prior on the background rates of infection

The joint prior of $(\boldsymbol{\alpha}, \lambda)$ (see Eq. (4)) is given by

$$\pi(\boldsymbol{\alpha}, \lambda) \propto \pi_\lambda(\lambda; b_\lambda) \prod_{g=1}^G \pi_\alpha(\alpha_g | \lambda). \quad (6)$$

In order to improve mixing, we used a non-centred parameterisation $\alpha_g^* = h(\alpha_g) = \frac{1}{\lambda}\alpha_g$, $\forall g$. Note that

$$\begin{aligned}\pi_{\alpha^*}(\alpha_g^* | \lambda) &= \pi_\alpha(h^{-1}(\alpha_g^*) | \lambda) \left| \frac{\partial \alpha_g}{\partial \alpha_g^*} \right| \\ &= \frac{1}{\lambda} e^{-\frac{1}{\lambda} \lambda \alpha_g^*} |\lambda| \\ &= e^{-\alpha_g^*}.\end{aligned}$$

Therefore, $\alpha_g^* \sim \text{Exp}(1)$, and Eq. (6), after non-centred parameterisation, will be given by:

$$\begin{aligned}\pi_{\alpha^*}(\boldsymbol{\alpha}^*, \lambda) &= \pi_\lambda(\lambda; b_\lambda) \prod_{g=1}^G \pi_{\alpha^*}(\alpha_g^* | \lambda) \\ &= b_\lambda e^{-b_\lambda \lambda} \prod_{g=1}^G e^{-\alpha_g^*}.\end{aligned}$$

Using log-transformations $\tilde{\alpha}_g^* = \log(\alpha_g^*)$, $\forall g$, and $\tilde{\lambda} = \log(\lambda)$, the joint prior of $(\tilde{\alpha}_g^*, \tilde{\lambda})$ (i.e., after log-transformation) is given by

$$\begin{aligned}\pi_{\tilde{\alpha}^*, \tilde{\lambda}}(\tilde{\alpha}_g^*, \tilde{\lambda}) &\propto e^{-b_\lambda e^{\tilde{\lambda}}} \left| \frac{\partial \lambda}{\partial \tilde{\lambda}} \right| \prod_{g=1}^G e^{-e^{\tilde{\alpha}_g^*}} \left| \frac{\partial \alpha_g^*}{\partial \tilde{\alpha}_g^*} \right| \\ &\propto e^{-b_\lambda e^{\tilde{\lambda}}} e^{\tilde{\lambda}} \prod_{g=1}^G e^{-e^{\tilde{\alpha}_g^*}} e^{\tilde{\alpha}_g^*},\end{aligned}$$

whose logarithm is

$$\log \left\{ \pi_{\tilde{\alpha}^*, \tilde{\lambda}}(\tilde{\alpha}_g^*, \tilde{\lambda}) \right\} \propto \tilde{\lambda} - b_\lambda e^{\tilde{\lambda}} + \sum_{g=1}^G \left(\tilde{\alpha}_g^* - e^{\tilde{\alpha}_g^*} \right).$$

The gradients used in HMC are then given by

$$\begin{aligned}\frac{\partial \log \left\{ \pi_{\tilde{\alpha}^*, \tilde{\lambda}}(\tilde{\alpha}_g^*, \tilde{\lambda}) \right\}}{\partial \tilde{\alpha}_g^*} &= 1 - e^{\tilde{\alpha}_g^*}, \quad g = 1, \dots, G, \\ \frac{\partial \log \left\{ \pi_{\tilde{\alpha}^*, \tilde{\lambda}}(\tilde{\alpha}_g^*, \tilde{\lambda}) \right\}}{\partial \tilde{\lambda}} &= 1 - b_\lambda e^{\tilde{\lambda}}.\end{aligned}$$

D1.1 Updating $\boldsymbol{\alpha}$, λ , β , q , τ , a , b , c

We used the log-transformations on $\boldsymbol{\alpha}$, λ , β , τ , a , b , and c as they are on $(0, \infty)$, and logistic transformation on q , which is on $(0, 1)$. The prior distributions chosen for these parameters are described in Section B3.

In the equations below, to simplify notation, we define

$$F_{g_{i,t-1}} := e^{\tilde{\alpha}_{g_i(t-1)}^*} e^{\tilde{\lambda}} + e^{\tilde{\beta}} K^q \frac{|\mathcal{I}_{i,t-1}|}{|\mathcal{M}_{i,t-1}|^q},$$

where $q = \frac{e^{\tilde{q}}}{1 + e^{\tilde{q}}}$. Both q and \tilde{q} terms may appear in the same equation for readability.

If the individual was born before its first social group started being monitored ($t_i^B < t_{0,i}$), then the likelihood function includes the probability of having survived up to time $t_{0,i}$, i.e. $S(t_{0,i} - t_i^B)$, where $t_{0,i} - t_i^B$ is the age of individual i at time $t_{0,i}$.

Time T_i is the end of the monitoring period of the last social group to which individual i belonged. If the last capture of individual i (denoted by T_i^{LC}) occurred after the monitoring period, this information is taken into account in the likelihood for the inference of death rates.

The joint conditional log-posterior of $(\tilde{\alpha}, \tilde{\lambda}, \tilde{\beta}, \tilde{q}, \tilde{\tau}, \tilde{a}, \tilde{b}, \tilde{c})$ is given by:

$$\begin{aligned}
\log \left(\pi(\tilde{\alpha}, \tilde{\lambda}, \tilde{\beta}, \tilde{q}, \tilde{\tau}, \tilde{a}, \tilde{b}, \tilde{c} \mid \cdot) \right) \propto & \left(\sum_{g=1}^G \left(\tilde{\alpha}_g^* - e^{\tilde{\alpha}_g^*} \right) \right) + \tilde{\lambda} - b_\lambda e^{\tilde{\lambda}} + a_\beta \tilde{\beta} - b_\beta e^{\tilde{\beta}} + \\
& a_q - (a_q + b_q) \log(1 + e^{\tilde{q}}) + a_\tau \tilde{\tau} - b_\tau e^{\tilde{\tau}} + a_a \tilde{a} - b_a e^{\tilde{a}} + a_b \tilde{b} - b_b e^{\tilde{b}} + \\
& a_c \tilde{c} - b_c e^{\tilde{c}} + \left(\sum_{i: t_i^B < t_{0,i}} \log S(t_{0,i} - t_i^B) \right) + \\
& \sum_{i=1}^m \sum_{t=t_{0,i}+1}^{T_i} \left[I_t^{[i]}(S, S) \left(\log(1 - \psi_t^{[i]}) - F_{g_{i,t-1}} \right) + \right. \\
& \quad I_t^{[i]}(S, E) \left(\log(1 - \psi_t^{[i]}) + \log \left(1 - \exp \left\{ -F_{g_{i,t-1}} \right\} \right) \right) + \\
& \quad I_t^{[i]}(E, E) \left(\log(1 - \psi_t^{[i]}) - 1/e^{\tilde{\tau}} \right) + \\
& \quad I_t^{[i]}(E, I) \left(\log(1 - \psi_t^{[i]}) + \log \left(1 - \exp \left\{ -1/e^{\tilde{\tau}} \right\} \right) \right) + \\
& \quad I_t^{[i]}(1, 1) \left(\log(1 - \psi_t^{[i]}) \right) + \\
& \quad \left. \left(I_t^{[i]}(S, D) + I_t^{[i]}(E, D) + I_t^{[i]}(I, D) \right) \log \left(\psi_t^{[i]} \right) \right] + \\
& \sum_{i: T_i^{\text{LC}} > T_i} \sum_{t=T_i+1}^{T_i^{\text{LC}}} \log \left(1 - \psi_t^{[i]} \right)
\end{aligned}$$

The gradients used in HMC are then given by

$$\begin{aligned}
\frac{\partial \log(\pi(\tilde{\alpha}_g^* \mid \cdot))}{\partial \tilde{\alpha}_g^*} = & 1 - e^{\tilde{\alpha}_g^*} + \\
& \sum_{i=1}^m \sum_{t=t_{0,i}+1}^{T_i} \left[-I_t^{[i]}(S, S) e^{\tilde{\alpha}_{g_{i(t-1)}}^*} e^{\tilde{\lambda}} + I_t^{[i]}(S, E) e^{\tilde{\alpha}_{g_{i(t-1)}}^*} e^{\tilde{\lambda}} \frac{\exp \left\{ -F_{g_{i,t-1}} \right\}}{1 - \exp \left\{ -F_{g_{i,t-1}} \right\}} \right],
\end{aligned}$$

for whenever $g_{i(t-1)} = g$, $g = 1, \dots, G$,

$$\begin{aligned}
\frac{\partial \log(\pi(\tilde{\lambda} | \cdot))}{\partial \tilde{\lambda}} &= 1 - b_\lambda e^{\tilde{\lambda}} + \\
&\sum_{i=1}^m \sum_{t=t_{0,i}+1}^{T_i} \left[-I_t^{[i]}(S, S) e^{\tilde{\alpha}_{g_i(t-1)}^* e^{\tilde{\lambda}}} + I_t^{[i]}(S, E) e^{\tilde{\alpha}_{g_i(t-1)}^* e^{\tilde{\lambda}}} \frac{\exp\{-F_{g_i,t-1}\}}{1 - \exp\{-F_{g_i,t-1}\}} \right], \\
\frac{\partial \log(\pi(\tilde{\beta} | \cdot))}{\partial \tilde{\beta}} &= a_\beta - b_\beta e^{\tilde{\beta}} + \\
&\sum_{i=1}^m \sum_{t=t_{0,i}+1}^{T_i} \left[-I_t^{[i]}(S, S) e^{\tilde{\beta}} \left(K^q \frac{|\mathcal{I}_{i,t-1}|}{|\mathcal{M}_{i,t-1}|^q} \right) + \right. \\
&\quad \left. I_t^{[i]}(S, E) e^{\tilde{\beta}} \left(K^q \frac{|\mathcal{I}_{i,t-1}|}{|\mathcal{M}_{i,t-1}|^q} \right) \frac{\exp\{-F_{g_i,t-1}\}}{1 - \exp\{-F_{g_i,t-1}\}} \right], \\
\frac{\partial \log(\pi(\tilde{q} | \cdot))}{\partial \tilde{q}} &= a_q - (a_q + b_q) \left(\frac{e^{\tilde{q}}}{1 + e^{\tilde{q}}} \right) + \\
&\sum_{i=1}^m \sum_{t=t_{0,i}+1}^{T_i} \left[I_t^{[i]}(S, S) e^{\tilde{\beta}} \log |\mathcal{M}_{i,t-1}| \left(K^q \frac{|\mathcal{I}_{i,t-1}|}{|\mathcal{M}_{i,t-1}|^q} \right) \frac{e^{\tilde{q}}}{(1 + e^{\tilde{q}})^2} + \right. \\
&\quad \left. - I_t^{[i]}(S, E) \frac{\exp\{-F_{g_i,t-1}\}}{1 - \exp\{-F_{g_i,t-1}\}} \times \right. \\
&\quad \left. e^{\tilde{\beta}} \log |\mathcal{M}_{i,t-1}| \left(K^q \frac{|\mathcal{I}_{i,t-1}|}{|\mathcal{M}_{i,t-1}|^q} \right) \frac{e^{\tilde{q}}}{(1 + e^{\tilde{q}})^2} \right], \\
\frac{\partial \log(\pi(\tilde{\tau} | \cdot))}{\partial \tilde{\tau}} &= a_\tau - b_\tau e^{\tilde{\tau}} + \\
&\sum_{i=1}^m \sum_{t=t_{0,i}+1}^{T_i} \left[I_t^{[i]}(E, E) 1/e^{\tilde{\tau}} - I_t^{[i]}(E, I) (1/e^{\tilde{\tau}}) \frac{\exp\{-1/e^{\tilde{\tau}}\}}{1 - \exp\{-1/e^{\tilde{\tau}}\}} \right].
\end{aligned}$$

In order to obtain the gradients of the survival distribution parameters, we need to know the derivatives of the survival function with respect to its parameters.

Using log-transformation on the survival distribution parameters, Eq. (2) becomes

$$\log S(u) = -e^{\tilde{c}}u + \frac{e^{\tilde{a}}}{e^{\tilde{b}}}(1 - e^{e^{\tilde{b}}u}),$$

and therefore

$$\begin{aligned}\frac{\partial \log S(u)}{\partial \tilde{a}} &= \frac{e^{\tilde{a}}}{e^{\tilde{b}}}(1 - e^{e^{\tilde{b}}u}), \\ \frac{\partial \log S(u)}{\partial \tilde{b}} &= -\frac{e^{\tilde{a}}}{e^{\tilde{b}}}(1 - e^{e^{\tilde{b}}u}) - \frac{e^{\tilde{a}}}{e^{\tilde{b}}}e^{e^{\tilde{b}}u}e^{\tilde{b}}u, \\ \frac{\partial \log S(u)}{\partial \tilde{c}} &= -e^{\tilde{c}}u.\end{aligned}$$

Let $p_t^{[i]} := 1 - \psi_t^{[i]}$ (see Eq. (1)) and let $u := \text{age}_t^{[i]}$. Thus, using the log-transformation on the parameters,

$$\log p_t^{[i]} = -e^{\tilde{c}} + \frac{e^{\tilde{a}}}{e^{\tilde{b}}}(e^{e^{\tilde{b}}(u-1)} - e^{e^{\tilde{b}}u}),$$

and thus

$$\begin{aligned}\frac{\partial \log p_t^{[i]}}{\partial \tilde{c}} &= -e^{\tilde{c}}, \\ \frac{\partial \log p_t^{[i]}}{\partial \tilde{a}} &= \frac{e^{\tilde{a}}}{e^{\tilde{b}}}(e^{e^{\tilde{b}}(u-1)} - e^{e^{\tilde{b}}u}), \\ \frac{\partial \log p_t^{[i]}}{\partial \tilde{b}} &= -\frac{e^{\tilde{a}}}{e^{\tilde{b}}}(e^{e^{\tilde{b}}(u-1)} - e^{e^{\tilde{b}}u}) + e^{\tilde{a}}((t-1)e^{e^{\tilde{b}}(t-1)} - te^{e^{\tilde{b}}t}).\end{aligned}$$

Finally, the derivative of $\log \psi_t^{[i]}$ with respect to $\tilde{x} \in (\tilde{a}, \tilde{b}, \tilde{c})$ is given by

$$\frac{\partial \log \psi_t^{[i]}}{\partial \tilde{a}} = \frac{1}{\psi_t^{[i]}} \frac{\partial \psi_t^{[i]}}{\partial \tilde{a}} = -\frac{1}{\psi_t^{[i]}} \frac{\partial p_t^{[i]}}{\partial \tilde{a}} = -\frac{1}{\psi_t^{[i]}} p_t^{[i]} \frac{\partial \log p_t^{[i]}}{\partial \tilde{a}}.$$

Therefore,

$$\begin{aligned}
\frac{\partial \log(\pi(\tilde{a} | \cdot))}{\partial \tilde{a}} &= a_a - b_a e^{\tilde{a}} + \sum_{i: t_i^B < t_{0,i}} \frac{\partial \log S(t_{0,i} - t_i^B)}{\partial \tilde{a}} + \\
&\quad \sum_{i=1}^m \sum_{t=t_{0,i}+1}^{T_i} \left[\left(I_t^{[i]}(S, S) + I_t^{[i]}(S, E) + I_t^{[i]}(E, E) + I_t^{[i]}(E, I) + I_t^{[i]}(I, I) \right) \frac{\partial \log(1 - \psi_t^{[i]})}{\partial \tilde{a}} + \right. \\
&\quad \left. \left(I_t^{[i]}(S, D) + I_t^{[i]}(E, D) + I_t^{[i]}(I, D) \right) \frac{\partial \log \psi_t^{[i]}}{\partial \tilde{a}} \right] + \\
&\quad \sum_{i: T_i^{\text{LC}} > T_i} \sum_{t=T_i+1}^{T_i^{\text{LC}}} \frac{\partial \log(1 - \psi_t^{[i]})}{\partial \tilde{a}}, \\
\frac{\partial \log(\pi(\tilde{b} | \cdot))}{\partial \tilde{b}} &= a_b - b_b e^{\tilde{b}} + \sum_{i: t_i^B < t_{0,i}} \frac{\partial \log S(t_{0,i} - t_i^B)}{\partial \tilde{b}} + \\
&\quad \sum_{i=1}^m \sum_{t=t_{0,i}+1}^{T_i} \left[\left(I_t^{[i]}(S, S) + I_t^{[i]}(S, E) + I_t^{[i]}(E, E) + I_t^{[i]}(E, I) + I_t^{[i]}(I, I) \right) \frac{\partial \log(1 - \psi_t^{[i]})}{\partial \tilde{b}} + \right. \\
&\quad \left. \left(I_t^{[i]}(S, D) + I_t^{[i]}(E, D) + I_t^{[i]}(I, D) \right) \frac{\partial \log \psi_t^{[i]}}{\partial \tilde{b}} \right] + \\
&\quad \sum_{i: T_i^{\text{LC}} > T_i} \sum_{t=T_i+1}^{T_i^{\text{LC}}} \frac{\partial \log(1 - \psi_t^{[i]})}{\partial \tilde{b}}, \\
\frac{\partial \log(\pi(\tilde{c} | \cdot))}{\partial \tilde{c}} &= a_c - b_c e^{\tilde{c}} + \sum_{i: t_i^B < t_{0,i}} \frac{\partial \log S(t_{0,i} - t_i^B)}{\partial \tilde{c}} + \\
&\quad \sum_{i=1}^m \sum_{t=t_{0,i}+1}^{T_i} \left[\left(I_t^{[i]}(S, S) + I_t^{[i]}(S, E) + I_t^{[i]}(E, E) + I_t^{[i]}(E, I) + I_t^{[i]}(I, I) \right) \frac{\partial \log(1 - \psi_t^{[i]})}{\partial \tilde{c}} + \right. \\
&\quad \left. \left(I_t^{[i]}(S, D) + I_t^{[i]}(E, D) + I_t^{[i]}(I, D) \right) \frac{\partial \log \psi_t^{[i]}}{\partial \tilde{c}} \right] + \\
&\quad \sum_{i: T_i^{\text{LC}} > T_i} \sum_{t=T_i+1}^{T_i^{\text{LC}}} \frac{\partial \log(1 - \psi_t^{[i]})}{\partial \tilde{c}}.
\end{aligned}$$

D1.2 Updating θ_j, ρ_j

Using logistic transformations and putting Beta priors on θ_j and ρ_j for each test j , the joint conditional log-posterior of $(\tilde{\theta}_j, \tilde{\rho}_j)$ is given by

$$\begin{aligned} \log(\pi(\tilde{\theta}_j, \tilde{\rho}_j | \dots)) &\propto a_{\tilde{\theta}}\tilde{\theta}_j - (a_{\tilde{\theta}} + b_{\tilde{\theta}})\log(1 + e^{\tilde{\theta}_j}) + a_{\tilde{\rho}}\tilde{\rho}_j - (a_{\tilde{\rho}} + b_{\tilde{\rho}})\log(1 + e^{\tilde{\rho}_j}) + \\ &\sum_{i=1}^m \sum_{o_{jt}^{[i]}=1} \left\{ \left[y_{jt}^{[i]} \log\left(\frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j}}\right) + (1 - y_{jt}^{[i]}) \log\left(1 - \frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j}}\right) \right] \mathbb{1}_{\{X_t^{[i]}=I\}} + \right. \\ &\quad \left[y_{jt}^{[i]} \left(\log\left(\frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j}}\right) + \log\left(\frac{e^{\tilde{\rho}_j}}{1 + e^{\tilde{\rho}_j}}\right) \right) + \right. \\ &\quad \left. \left. (1 - y_{jt}^{[i]}) \log\left(1 - \frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j}} \frac{e^{\tilde{\rho}_j}}{1 + e^{\tilde{\rho}_j}}\right) \right] \mathbb{1}_{\{X_t^{[i]}=E\}} \right\} \\ &\propto a_{\tilde{\theta}}\tilde{\theta}_j - (a_{\tilde{\theta}} + b_{\tilde{\theta}})\log(1 + e^{\tilde{\theta}_j}) + a_{\tilde{\rho}}\tilde{\rho}_j - (a_{\tilde{\rho}} + b_{\tilde{\rho}})\log(1 + e^{\tilde{\rho}_j}) + \\ &\sum_{i=1}^m \sum_{o_{jt}^{[i]}=1} \left\{ \left[y_{jt}^{[i]} (\tilde{\theta}_j - \log(1 + e^{\tilde{\theta}_j})) - (1 - y_{jt}^{[i]}) \log(1 + e^{\tilde{\theta}_j}) \right] \mathbb{1}_{\{X_t^{[i]}=I\}} + \right. \\ &\quad \left[y_{jt}^{[i]} \left(\tilde{\theta}_j - \log(1 + e^{\tilde{\theta}_j}) + \tilde{\rho}_j - \log(1 + e^{\tilde{\rho}_j}) \right) + \right. \\ &\quad \left. \left. (1 - y_{jt}^{[i]}) \left(\log(1 + e^{\tilde{\theta}_j} + e^{\tilde{\rho}_j}) - \log(1 + e^{\tilde{\theta}_j}) - \log(1 + e^{\tilde{\rho}_j}) \right) \right] \mathbb{1}_{\{X_t^{[i]}=E\}} \right\}. \end{aligned}$$

The gradients used in HMC are therefore given by

$$\begin{aligned} \frac{\partial \log(\pi(\tilde{\theta}_j | \dots))}{\partial \tilde{\theta}_j} &= a_{\theta} - (a_{\theta} + b_{\theta}) \frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j}} + \\ &\sum_{i=1}^m \sum_{o_{jt}^{[i]}=1} \left\{ \left(y_{jt}^{[i]} \left(1 - \frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j}} \right) - (1 - y_{jt}^{[i]}) \frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j}} \right) \mathbb{1}_{\{X_t^{[i]}=I\}} + \right. \\ &\quad \left. \left(y_{jt}^{[i]} \left(1 - \frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j}} \right) + (1 - y_{jt}^{[i]}) \left(\frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j} + e^{\tilde{\rho}_j}} - \frac{e^{\tilde{\theta}_j}}{1 + e^{\tilde{\theta}_j}} \right) \right) \mathbb{1}_{\{X_t^{[i]}=E\}} \right\}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \log(\pi(\tilde{\rho}_j | \dots))}{\partial \tilde{\rho}_j} &= a_{\rho} - (a_{\rho} + b_{\rho}) \frac{e^{\tilde{\rho}_j}}{1 + e^{\tilde{\rho}_j}} + \\ &\sum_{i=1}^m \sum_{o_{jt}^{[i]}=1} \left\{ \left(y_{jt}^{[i]} \left(1 - \frac{e^{\tilde{\rho}_j}}{1 + e^{\tilde{\rho}_j}} \right) + (1 - y_{jt}^{[i]}) \left(\frac{e^{\tilde{\rho}_j}}{1 + e^{\tilde{\theta}_j} + e^{\tilde{\rho}_j}} - \frac{e^{\tilde{\rho}_j}}{1 + e^{\tilde{\rho}_j}} \right) \right) \mathbb{1}_{\{X_t^{[i]}=E\}} \right\}. \end{aligned}$$

D2 Gibbs sampling updates

In this section, we show the conditional distributions used in Gibbs sampling updates for the remaining parameters. Prior choices are described in Section B3.

D2.1 Updating ϕ_j

The conditional posterior distribution of ϕ_j is given by

$$\pi(\phi_j \mid \dots) \propto \phi_j^{a_{\phi_j}-1} (1-\phi_j)^{b_{\phi_j}-1} \prod_{i=1}^m \prod_{\substack{X_t^{[i]}=S; \\ o_{jt}^{[i]}=1}} \left[(1-\phi_j)^{y_{jt}^{[i]}} \phi_j^{1-y_{jt}^{[i]}} \right].$$

Therefore, in the Gibbs step,

$$\phi_j \sim \text{Beta}\left(a_{\phi_j} + \sum_{i=1}^m \sum_{\substack{X_t^{[i]}=S; \\ o_{jt}^{[i]}=1}} (1-y_{jt}^{[i]}), b_{\phi_j} + \sum_{i=1}^m \sum_{\substack{X_t^{[i]}=S; \\ o_{jt}^{[i]}=1}} y_{jt}^{[i]}\right).$$

D2.2 Updating η_ℓ

For conditional posterior distribution for the quarterly capture probability η_ℓ , $\ell = 1, \dots, 4$, is given by

$$\begin{aligned} \pi(\eta_\ell \mid \dots) &\propto \eta_\ell^{a_{\eta_\ell}-1} (1-\eta_\ell)^{b_{\eta_\ell}-1} \prod_{i=1}^m \prod_{\substack{t=1; \\ X_t^{[i]}\in\{S,E,I\}}} \mathbb{1}_{\{s_t=\ell\}} \left[\eta_\ell^{w_t^{[i]}} (1-\eta_\ell)^{1-w_t^{[i]}} \right] \\ &\propto \eta_\ell^{a_{\eta_\ell}-1+\sum_{i=1}^m \sum_{t=1; X_t^{[i]}\in\{S,E,I\}} \mathbb{1}_{\{s_t=\ell\}} w_t^{[i]}} (1-\eta_\ell)^{b_{\eta_\ell}-1+\sum_{i=1}^m \sum_{t=1; X_t^{[i]}\in\{S,E,I\}} \sum_{t=1}^T \mathbb{1}_{\{s_t=\ell\}} (1-w_t^{[i]})}. \end{aligned}$$

Therefore, the Gibbs step consists of drawing from

$$\eta_\ell \sim \text{Beta}\left(a_{\eta_\ell} + \sum_{i=1}^m \sum_{\substack{t=1; \\ X_t^{[i]}\in\{S,E,I\}}} \mathbb{1}_{\{s_t=\ell\}} w_t^{[i]}, b_{\eta_\ell} + \sum_{i=1}^m \sum_{\substack{t=1; \\ X_t^{[i]}\in\{S,E,I\}}} \mathbb{1}_{\{s_t=\ell\}} (1-w_t^{[i]})\right).$$

D2.3 Updating ν_S, ν_E, ν_I

The joint conditional posterior distribution of (ν_S, ν_E, ν_I) , by using a Dirichlet prior $\text{Dir}(1, 1, 1)$, is

$$\pi(\nu_S, \nu_E, \nu_I \mid \dots) \propto \prod_{i:t_i^B < t_{0,i}} \left[\nu_S^{\mathbb{1}_{\{X_{t_{0,i}}^{[i]}=S\}}} \nu_E^{\mathbb{1}_{\{X_{t_{0,i}}^{[i]}=E\}}} \nu_I^{\mathbb{1}_{\{X_{t_{0,i}}^{[i]}=I\}}} \right].$$

Note that we use only information from the individuals which were born before the monitoring period, i.e. $t_i^B < t_{0,i}$. Thus, values for (ν_S, ν_E, ν_I) are sampling from

$$(\nu_S, \nu_E, \nu_I) \sim \text{Dir}\left(1 + \sum_{i:t_i^B < t_{0,i}} \mathbb{1}_{\{X_{t_{0,i}}^{[i]}=S\}}, 1 + \sum_{i:t_i^B < t_{0,i}} \mathbb{1}_{\{X_{t_{0,i}}^{[i]}=E\}}, 1 + \sum_{i:t_i^B < t_{0,i}} \mathbb{1}_{\{X_{t_{0,i}}^{[i]}=I\}}\right).$$

D3 Metropolis-Hastings update for ξ

Brock changepoint ξ is the only parameter updated by Metropolis-Hastings. The conditional log-posterior of ξ is given by

$$\begin{aligned} \log(\pi(\xi | \dots)) \propto & -\frac{1}{2\sigma_\xi^2}(\xi - \mu_\xi)^2 + \\ & \sum_{i=1}^m \sum_{o_{jt}^{[i]}=1} \left\{ \left[y_{jt}^{[i]} \log(\theta_j) + (1 - y_{jt}^{[i]}) \log(1 - \theta_j) \right] \mathbb{1}_{\{X_t^{[i]}=I\}} + \right. \\ & \left. \left[y_{jt}^{[i]} \log(\rho_j \theta_j) + (1 - y_{jt}^{[i]}) \log(1 - \rho_j \theta_j) \right] \mathbb{1}_{\{X_t^{[i]}=E\}} + \right. \\ & \left. \left[y_{jt}^{[i]} \log(1 - \phi_j) + (1 - y_{jt}^{[i]}) \log(\phi_j) \right] \mathbb{1}_{\{X_t^{[i]}=S\}} \right\} \end{aligned}$$

where $o_{jt}^{[i]}$ and $y_{jt}^{[i]}$, for tests $j \in \{1, 2\}$, depend on ξ (see Eq. (3)) in the model with Brock changepoint.

E Individual effective reproduction number

[12] derive a likelihood-based estimator for the individual effective reproduction number, R_i , for an individual i as:

$$R_i = \sum_{k|t_k^E > t_i^E} p_{(k,i)}. \quad (7)$$

Here $p_{(k,i)}$ is the relative likelihood of individual k being infected by individual i , and hence R_i is the sum of these relative likelihoods over all possible future infections $\{k; t_k^E > t_i^E\}$. [12] derive an estimator for $p_{(k,i)}$ based on the generation-time distribution $w(\cdot)$, such that:

$$p_{(k,i)} = \frac{w(t_k^E - t_i^E)}{\sum_{j|t_k^E > t_j^E} w(t_k^E - t_j^E)}.$$

Ordinarily, since the specific who-infected-whom infection network is unknown it is necessary to integrate over all possible infection networks to generate an estimate of R_i , which is computationally intractable unless very few individuals are infected. However, under the assumption that the prior probability of any given infection network being true is constant, [12] show that (7) is equivalent to integrating over all possible networks, and is substantially more efficient.

Since our model is not defined in terms of the generation-time distribution, we instead derive a variant of this estimator that utilises the exact likelihoods that we have in our model. Furthermore, to use the [12] estimator directly also requires knowledge of the infection times, which are unknown. [12] used the difference between reported dates of symptom onset as a proxy for $t_k^E - t_i^E$. Our approach does not need any proxy, as the iFFBS algorithm generates random samples for the unobserved event times at each iteration of the MCMC algorithm, and we can condition on these directly to generate random samples from the posterior distribution of R_i . Taken over the whole MCMC chain, these can be used to estimate the posterior distribution for R_i which numerically integrates over the relevant hidden states. In addition, we adjust for the fact that we have additional sources of infectious pressure through the background rates-of-infection.

E1 Discrete event times

At a given iteration of the MCMC, we have imputed infection times, t_k^E , for N^E infected individuals ($k = 1, \dots, N^E$). In practice N^E may change between iterations, since it depends on the infection states of individuals, which are unknown and thus imputed using the iFFBS algorithm. For brevity, we ignore any dependence on the specific MCMC iteration in the discussion below.

Hence, at a given iteration we have a set of simulated discrete infection times, t_k^E . From [12] we can define:

$$R_i = \sum_{k:t_k^E > t_i^E} p_{(k,i)}, \quad (8)$$

where $p_{(k,i)}$ is the relative likelihood that individual i infects individual k . Using the same assumptions as [12], and assuming no background rate-of-infection terms, this reduces to:

$$p_{(k,i)} = \frac{P_{ki} (t_k^E - 1 < T_k^E < t_k^E \mid T_k^E > t_k^E - 1) \frac{1}{|\mathcal{I}_{k,t_k^E-1}|} \prod_{t=t_k^B+1}^{t_k^E-1} P_{ki} (T_k^E > t \mid T_k^E > t-1)}{\sum_{j \in \mathcal{I}_{k,t_k^E-1}} P_{kj} (t_k^E - 1 < T_k^E < t_k^E \mid T_k^E > t_k^E - 1) \frac{1}{|\mathcal{I}_{k,t_k^E-1}|} \prod_{t=t_k^B+1}^{t_k^E-1} P_{kj} (T_k^E > t \mid T_k^E > t-1)} \quad (9)$$

which simplifies to:

$$p_{(k,i)} = \begin{cases} \frac{1}{|\mathcal{I}_{k,t_k^E-1}|} & \text{if } i \in \mathcal{I}_{k,t_k^E-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Here, $P_{ki}(\cdot)$ correspond to probability statements regarding infection events for individual k due to the infectious pressure generated by individual i .

In (9), $\frac{1}{|\mathcal{I}_{k,t_k^E-1}|}$ refers to the probability that individual k was infected in $[t_k^E - 1, t_k^E]$ by one of the individuals which were infectious at $t_k^E - 1$ in the same social group, given that individual k was susceptible at $t_k^E - 1$. This comes from the fact that each individual $j \in \mathcal{I}_{k,t_k^E-1}$ is assumed to have the same probability of infecting k .

Adjusting for background rates-of-infection

For our model we also need to adjust the estimator for R_i to account for infectious pressure due to the background rate-of-infection terms. Similarly to Eq. (9) we can consider:

$$p_{(k,i)} \propto P_{ki} (t_k^E - 1 < T_k^E < t_k^E \mid T_k^E > t_k^E - 1) \frac{\beta K^q / |\mathcal{M}_{k,t_k^E-1}|^q}{\alpha_{g_k(t_k^E-1)} + |\mathcal{I}_{k,t_k^E-1}| \beta K^q / |\mathcal{M}_{k,t_k^E-1}|^q} \prod_{t=t_k^B+1}^{t_k^E-1} P_{ki} (T_k^E > t \mid T_k^E > t-1),$$

where the constant-of-proportionality is given by:

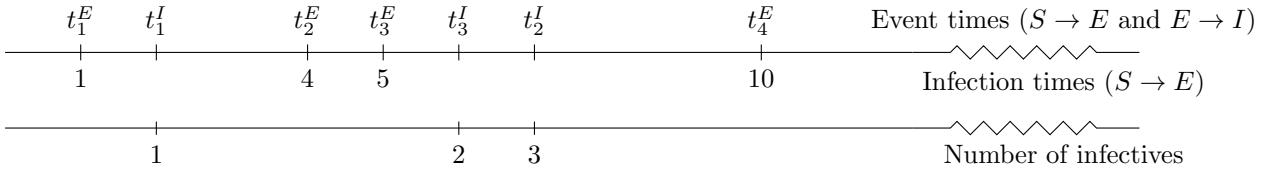
$$P_{k\alpha} (t_k^E - 1 < T_k^E < t_k^E \mid T_k^E > t_k^E - 1) \frac{\alpha_{g_k(t_k^E-1)}}{\alpha_{g_k(t_k^E-1)} + |\mathcal{I}_{k,t_k^E-1}| \beta K^q / |\mathcal{M}_{k,t_k^E-1}|^q} \prod_{t=t_k^B+1}^{t_k^E-1} P_{k\alpha} (T_k^E > t \mid T_k^E > t-1) + \sum_{j \in \mathcal{I}_{k,t_k^E-1}} \left[P_{ki} (t_k^E - 1 < T_k^E < t_k^E \mid T_k^E > t_k^E - 1) \frac{\beta K^q / |\mathcal{M}_{k,t_k^E-1}|^q}{\alpha_{g_k(t_k^E-1)} + |\mathcal{I}_{k,t_k^E-1}| \beta K^q / |\mathcal{M}_{k,t_k^E-1}|^q} \prod_{t=t_k^B+1}^{t_k^E-1} P_{ki} (T_k^E > t \mid T_k^E > t-1) \right].$$

Therefore,

$$p_{(k,i)} = \begin{cases} \frac{\beta K^q / |\mathcal{M}_{k,t_k^E-1}|^q}{\alpha_{g_k(t_k^E-1)} + |\mathcal{I}_{k,t_k^E-1}| \beta K^q / |\mathcal{M}_{k,t_k^E-1}|^q} & \text{if } i \in \mathcal{I}_{k,t_k^E-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Illustration

For simplicity, assume that all the m individuals remain alive and belong to the same social group g . Let t_i^E and t_i^I be the times at which individual i becomes exposed and infectious, respectively.



Let $p_{(k,i)}$ be the probability that badger i infects badger k at each new infection time t_k^E , and let $p_{(k,\alpha_g)}$ be the probability that badger k is infected via the background rate-of-transmission term. Below we illustrate how R_i is obtained for the individuals that became infectious ($i = 1, 2, 3$).

Newly infected badger (k)	1	2	3	4
Infection time (t_k^E)	1	4	5	10
Total rate (r_{t-1})	α_g	$\alpha_g + \beta \frac{K^q}{m^q}$	$\alpha_g + \beta \frac{K^q}{m^q}$	$\alpha_g + 3\beta \frac{K^q}{m^q}$
$p_{(k,1)}$	0	$\beta \frac{K^q}{m^q} / r_3$	$\beta \frac{K^q}{m^q} / r_4$	$\beta \frac{K^q}{m^q} / r_9$
$p_{(k,2)}$	0	0	0	$\beta \frac{K^q}{m^q} / r_9$
$p_{(k,3)}$	0	0	0	$\beta \frac{K^q}{m^q} / r_9$
$p_{(k,\alpha_g)}$	1	α_g / r_3	α_g / r_4	α_g / r_9

We define the individual reproduction number, R_i from (8) as:

$$R_1 = \beta \frac{K^q}{m^q} \left(\frac{1}{r_3} + \frac{1}{r_4} + \frac{1}{r_9} \right),$$

$$R_2 = \beta \frac{K^q}{m^q} \left(\frac{1}{r_9} \right),$$

$$R_3 = \beta \frac{K^q}{m^q} \left(\frac{1}{r_9} \right).$$

The estimate of the posterior distribution for each R_i is obtained from the values obtained for each iteration of the MCMC. The population reproduction number is defined as the arithmetic mean of R_i over all badgers.

We note that it is in theory possible to derive an effective R_{α_g} here for the background rate-of-infection terms, but since the background rates of infection are constant in this model, then R_{α_g} is not constrained by any kind of infectious period, and so from (8) we can see that R_{α_g} would simply increase as the total number of infected badgers increases, so it does not make much biological sense to do this here.

E2 Population-level effective reproduction numbers

There are various ways in which population-level effective reproduction numbers can be defined, see for example [12, 7, 5]. Here we do something more akin to the approach of [12], in which we retrospectively generate the expected number of cases infected by an animal across its lifetime. To obtain a population-level estimate, we simply average across individuals, giving:

$$R_t = \frac{1}{N_t^D} \sum_{i=1}^{N_t^D} R_i,$$

where R_i is defined in (8) and N_t^D is the number of badgers that have died up to time t . Hence R_t is *retrospective*, in the sense that it is an estimate of the expected number of secondary infections that were caused by a random infected individual across its whole lifetime. We restrict to only looking at individuals who have died in order to keep within the paradigm of trying to understand what has happened in the population so far, rather than what will happen going forwards. For the latter we could perform forwards-simulations from the model under some assumptions about future movements of animals, that would have to be determined.

F Relative rates of infection

For each iteration of the MCMC, the relative contribution of direct badger-to-badger transmission on any infected badger k at its infection time t_k^E can be calculated as:

$$\frac{\beta K^q |\mathcal{I}_{k,t_k^E-1}| / |\mathcal{M}_{k,t_k^E-1}|^q}{\alpha_{g_k(t_k^E-1)} + \beta K^q |\mathcal{I}_{k,t_k^E-1}| / |\mathcal{M}_{k,t_k^E-1}|^q}.$$

These can then be aggregated at any spatio-temporal resolution, and repeating this process for different MCMC samples allows for a predictive distribution to be built empirically.

G Additional results

G1 Mixing

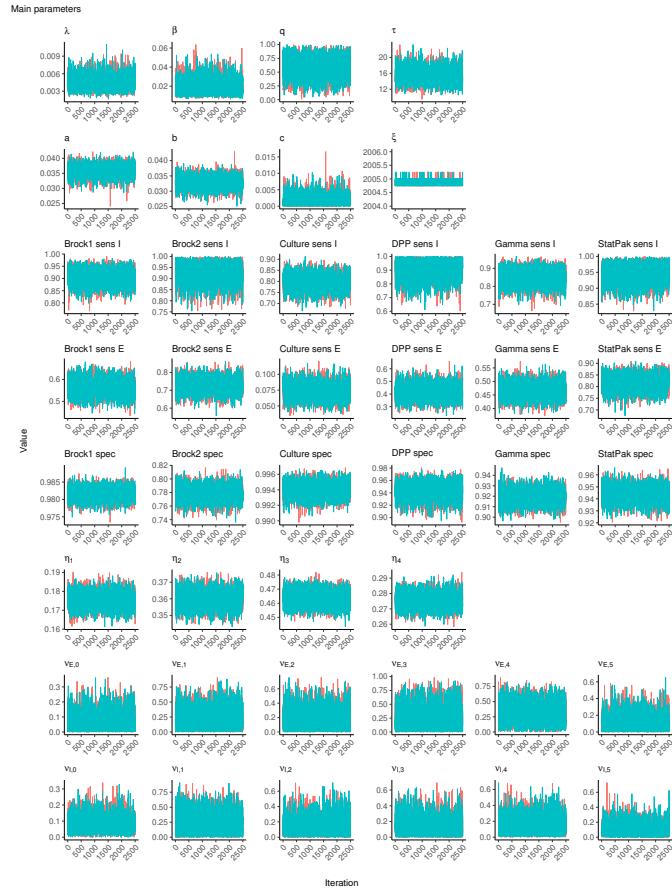


Figure B: Traceplots of all parameters (except background rates of infection).

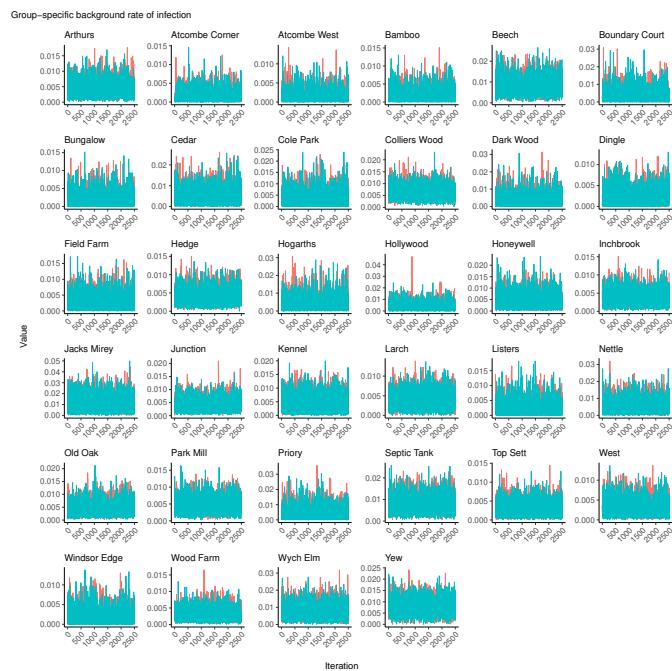


Figure C: Traceplots of group-specific background rates of infection.

G2 Additional model outputs

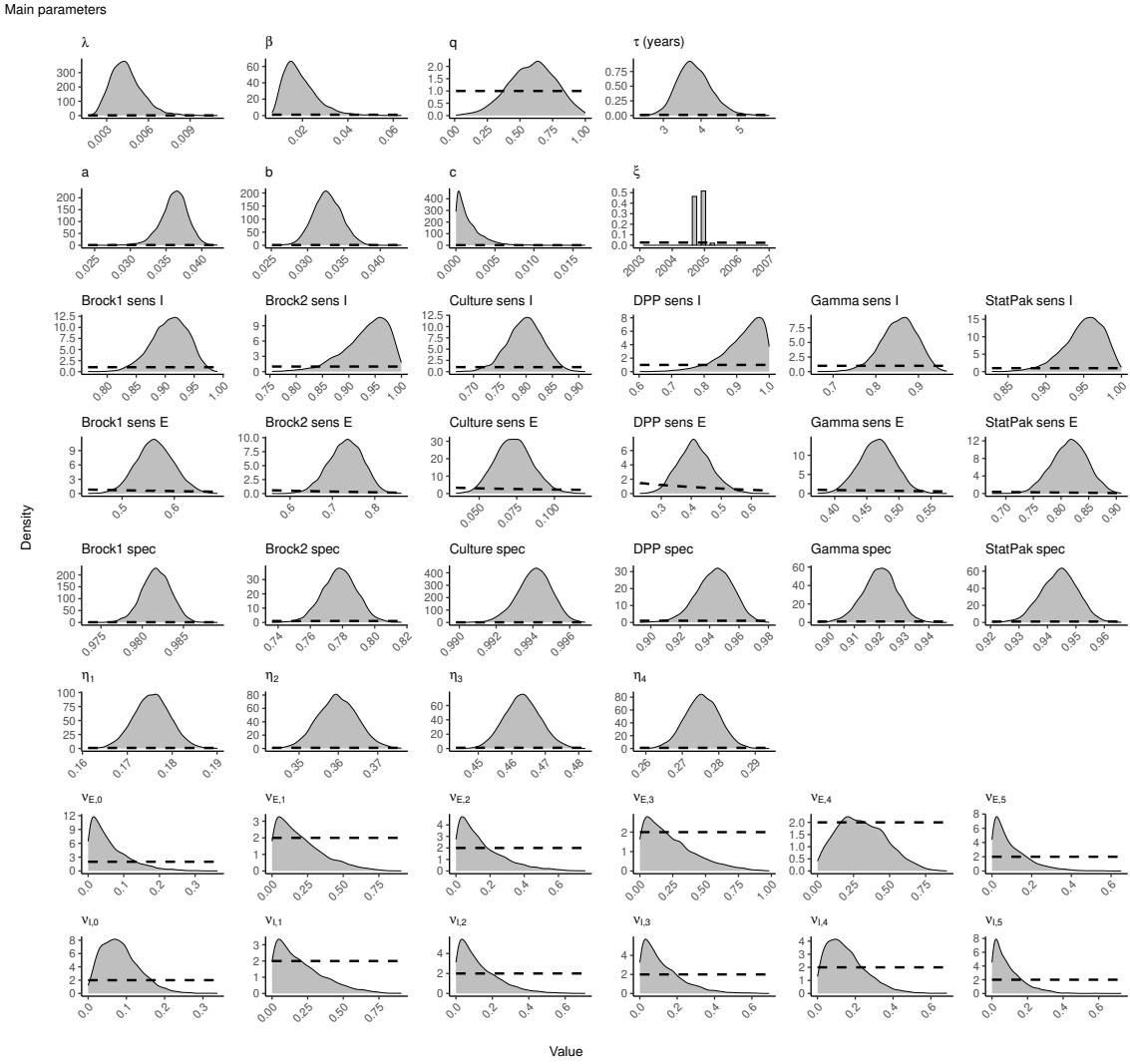


Figure D: Posterior densities for all parameters except the background rates of infection. Dashed lines represent the prior densities.

Group-specific background rate of infection

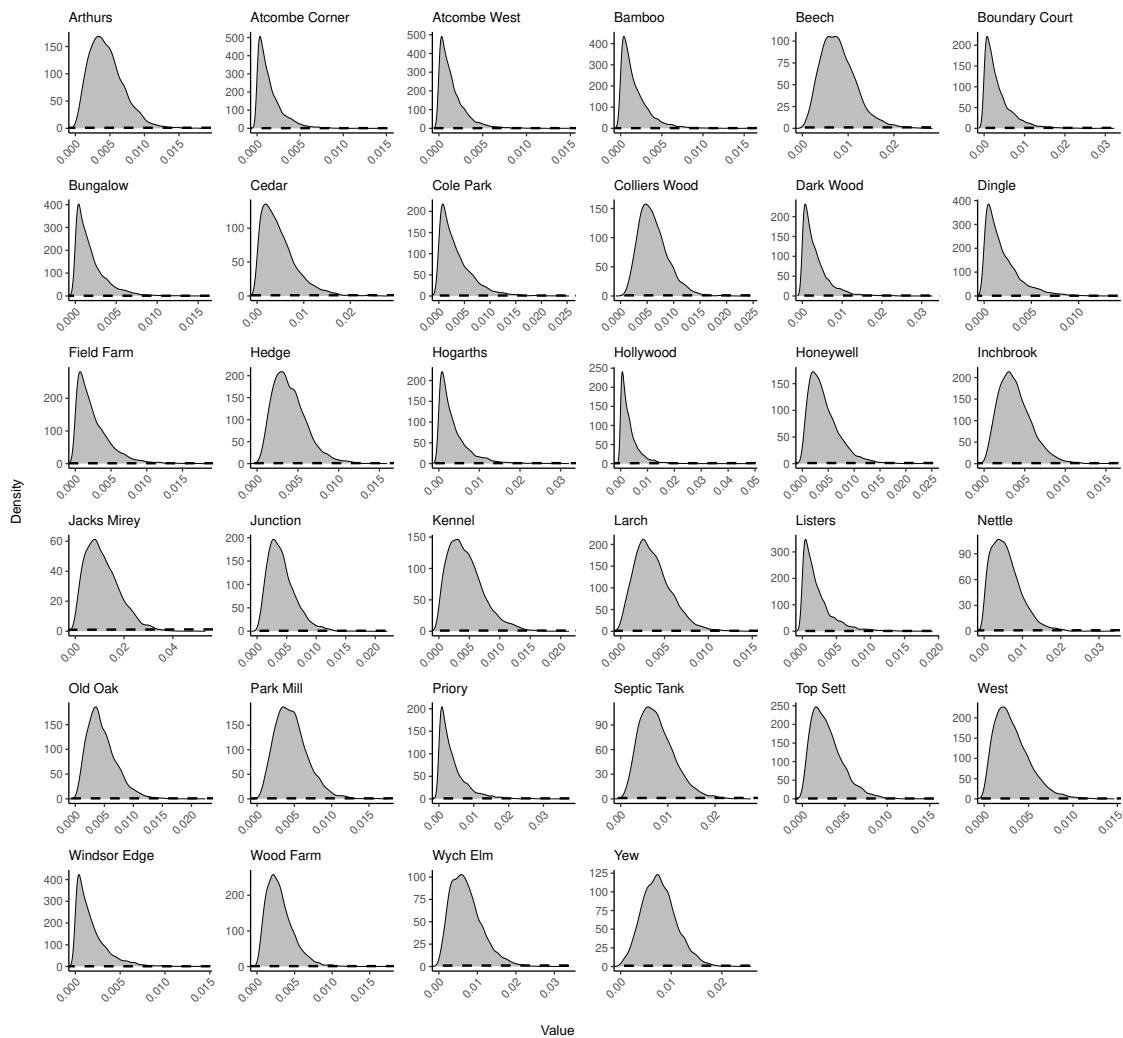
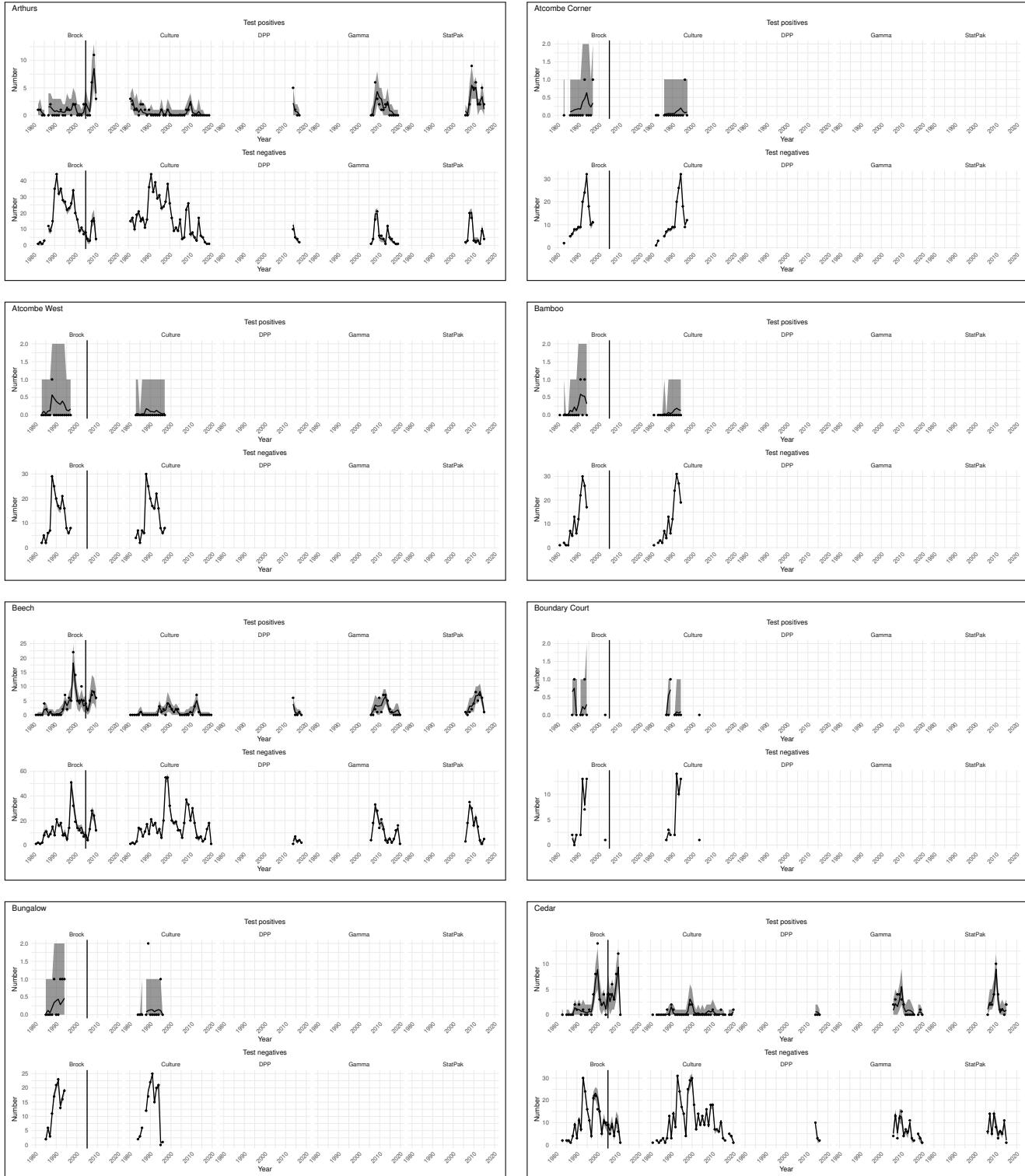


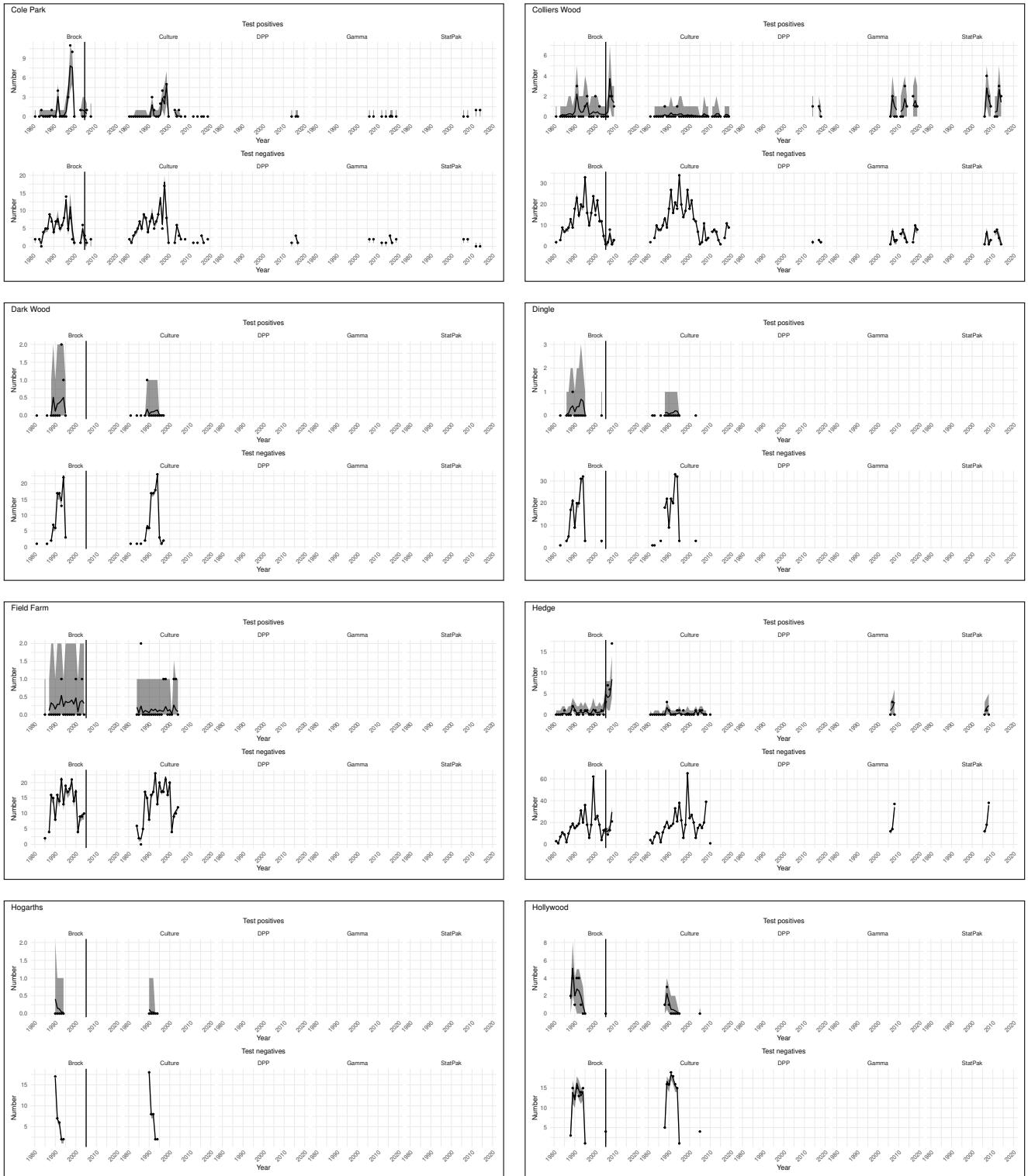
Figure E: Posterior densities of the background rates of infection. Dashed lines represent the prior densities.

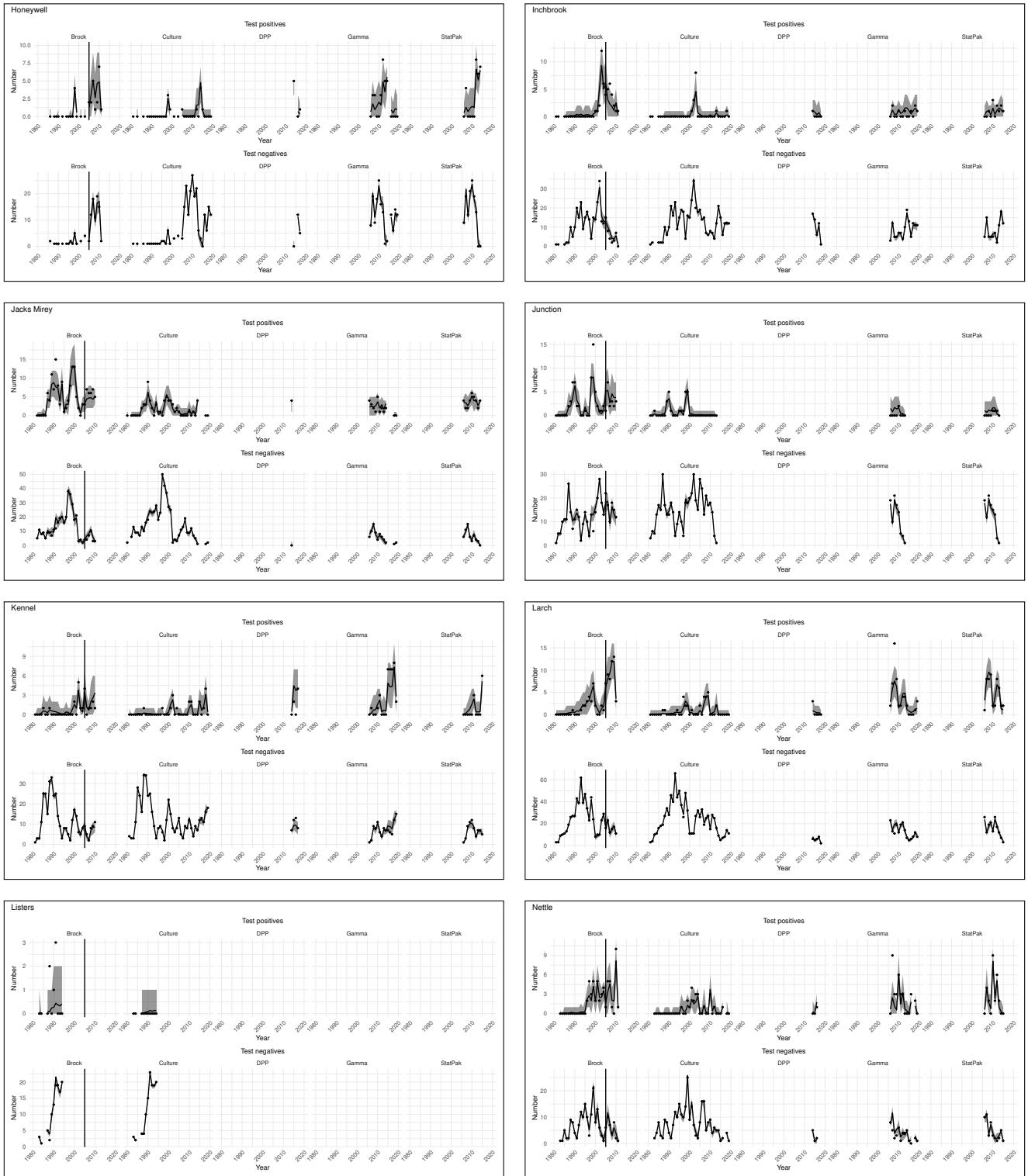
Group	Mean	95% CI
Arthurs	0.0045	(0.00088, 0.01)
Atcombe Corner	0.0015	(0.000036, 0.0054)
Atcombe West	0.0015	(0.000037, 0.0054)
Bamboo	0.0017	(0.000049, 0.0062)
Beech	0.008	(0.002, 0.017)
Boundary Court	0.0034	(0.000082, 0.013)
Bungalow	0.0019	(0.000046, 0.0068)
Cedar	0.0047	(0.00027, 0.014)
Cole Park	0.0034	(0.000084, 0.012)
Colliers Wood	0.0062	(0.002, 0.013)
Dark Wood	0.0033	(0.000085, 0.012)
Dingle	0.0019	(0.000051, 0.0068)
Field Farm	0.0025	(0.000077, 0.0086)
Hedge	0.004	(0.001, 0.0089)
Hogarths	0.0036	(0.000081, 0.014)
Hollywood	0.0032	(0.000063, 0.013)
Honeywell	0.0042	(0.00049, 0.012)
Inchbrook	0.0038	(0.00081, 0.0085)
Jacks Mirey	0.011	(0.0012, 0.028)
Junction	0.0042	(0.00087, 0.0098)
Kennel	0.0044	(0.00036, 0.012)
Larch	0.0037	(0.00055, 0.0084)
Listers	0.0022	(0.000061, 0.0084)
Nettle	0.0058	(0.00039, 0.015)
Old Oak	0.0046	(0.001, 0.011)
Park Mill	0.0046	(0.0012, 0.0095)
Priory	0.0039	(0.00011, 0.015)
Septic Tank	0.0077	(0.0021, 0.016)
Top Sett	0.0029	(0.00043, 0.0074)
West	0.0032	(0.00046, 0.008)
Windsor Edge	0.0018	(0.000048, 0.0066)
Wood Farm	0.003	(0.00058, 0.0073)
Wych Elm	0.0072	(0.0012, 0.017)
Yew	0.0077	(0.0019, 0.015)

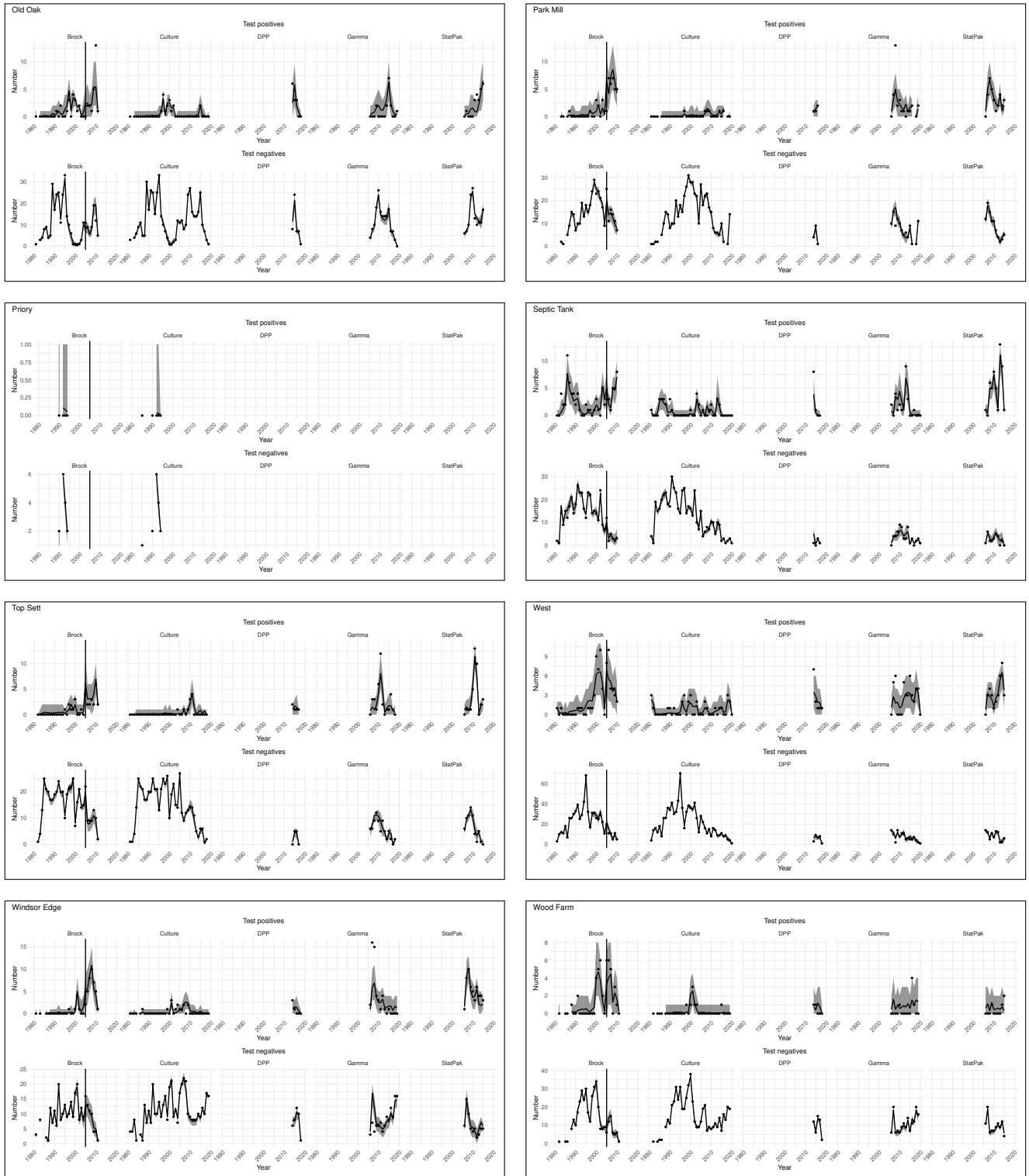
Table A: Posterior means and 95% credible intervals for group-specific background rate-of-infection terms. Shown to 2 significant figures.

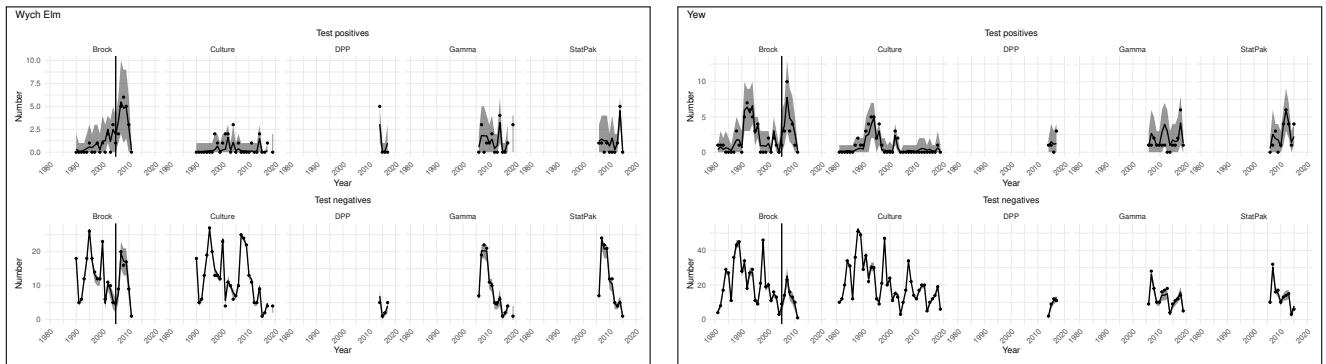
G3 Social group model fits











G4 Results for model without Brock changepoint

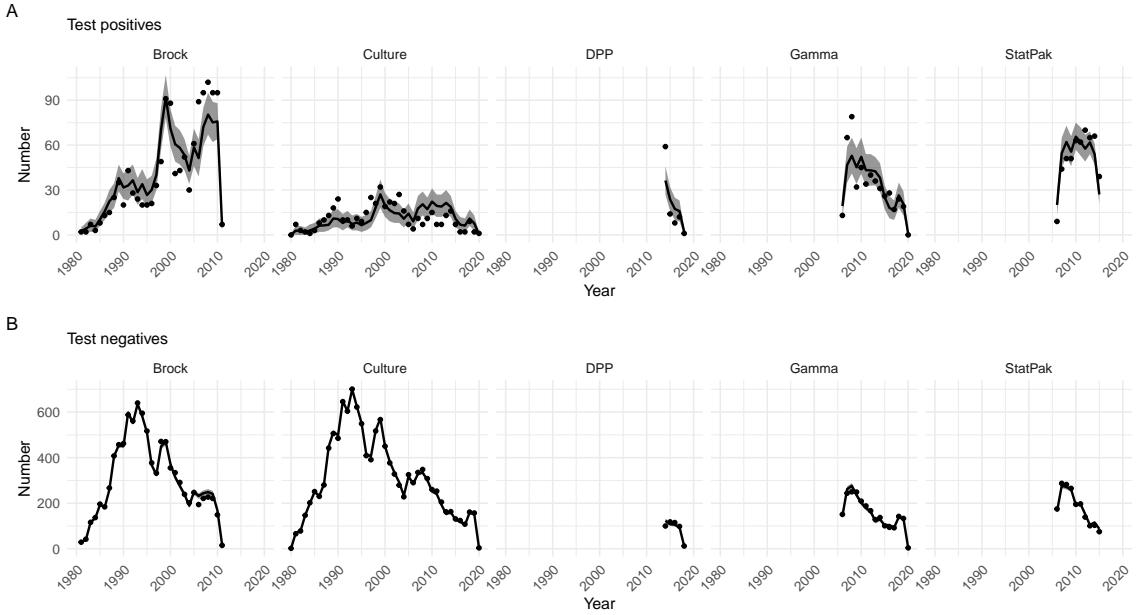


Figure F: Posterior predictive distribution for the number of (A) test-positive and (B) test-negative results in each year against the observed data. Posterior means are indicated by black lines and 95% credible intervals by grey ribbons. Data are represented by black points.

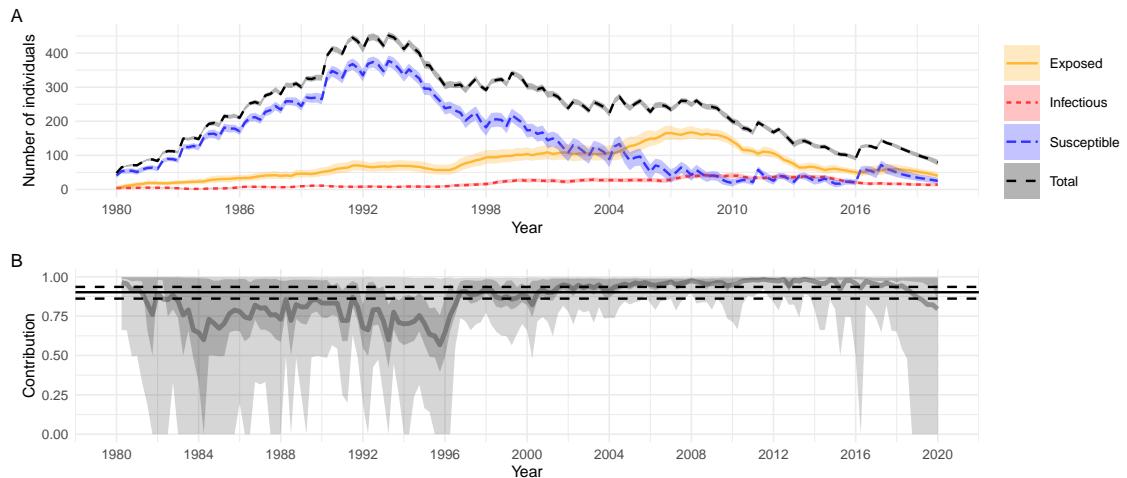


Figure G: Posterior predictive summaries: (A) posterior predictive means and 95% prediction intervals for the number of individuals in each epidemiological state over time; (B) posterior predictive means and credible intervals (50% and 95%) for the relative contribution of badger-to-badger transmission vs. background rates-of-infection on new infections over time (the solid black line is the average contribution over all time points, with the dashed lines giving the corresponding 95% credible interval).

Type	Parameter	Mean	95% CI
Transmission	β	0.092	(0.041, 0.2)
	q	0.64	(0.19, 0.97)
	τ (years)	7	(5.9, 8.5)
Mortality	a	0.036	(0.032, 0.039)
	b	0.033	(0.029, 0.037)
	c	0.0016	(0.000036, 0.0058)
Sensitivity (E)	Brock	0.17	(0.14, 0.19)
	Culture	0.015	(0.011, 0.02)
	DPP	0.072	(0.039, 0.11)
	Gamma	0.11	(0.096, 0.13)
	StatPak	0.078	(0.063, 0.095)
Sensitivity (I)	Brock	0.88	(0.84, 0.91)
	Culture	0.32	(0.29, 0.35)
	DPP	0.63	(0.53, 0.74)
	Gamma	0.57	(0.53, 0.62)
	StatPak	0.93	(0.89, 0.96)
Specificity	Brock	0.99	(0.99, 0.99)
	Culture	1	(0.99, 1)
	DPP	0.95	(0.88, 1)
	Gamma	0.99	(0.98, 1)
	StatPak	0.97	(0.92, 1)
Capture Probability	Winter	0.18	(0.17, 0.18)
	Spring	0.36	(0.35, 0.37)
	Summer	0.46	(0.45, 0.47)
	Autumn	0.27	(0.27, 0.28)

Table B: Posterior means and 95% credible intervals for key model parameters (to 2 significant figures).

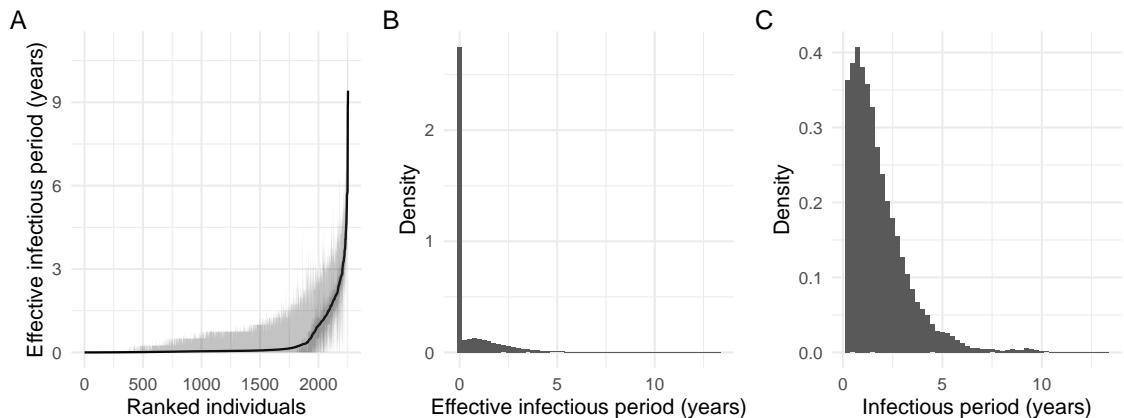


Figure H: (A) Posterior means and credible intervals (50% and 95%) for the individual effective infectious period distribution across all individuals, ranked in increasing order of posterior means; (B) posterior distribution for the population-averaged *effective* infectious period distribution; (C) posterior distribution for the population-averaged infectious period distribution.

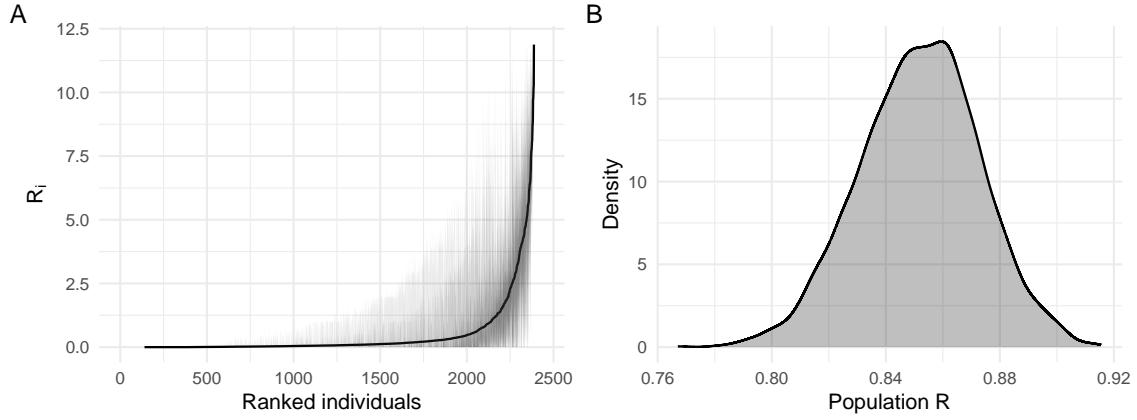


Figure I: (A) Posterior means and credible intervals (50% and 95%) for the individual reproduction number for all individuals, ranked in increasing order of posterior means, (B) posterior distribution for the population reproduction number.

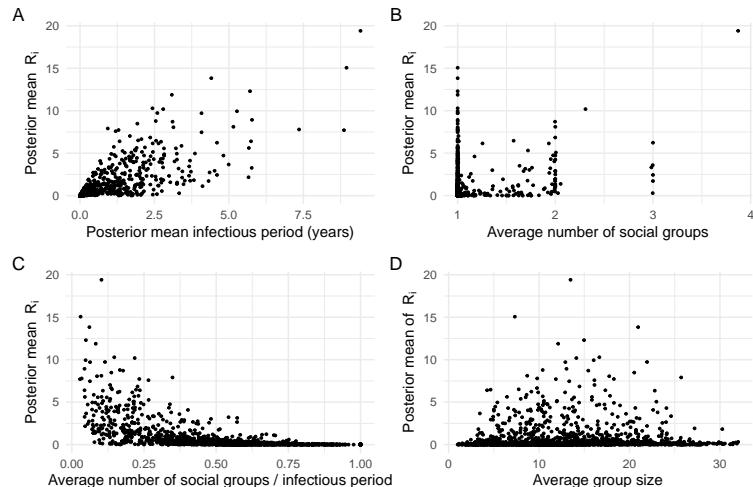


Figure J: Posterior mean estimates of the individual reproduction numbers against (A) posterior mean duration of the infectious period, (B) average number of social groups an individual belonged to, (C) average number of social groups an individual belonged to divided by the length of the infectious period, and (D) the average group size for social groups that the individual belonged to.

References

- [1] Matthew Brand. Coupled hidden Markov models for modeling interacting processes. Technical Report 405, The Media Lab, Massachusetts Institute of Technology, 1997.
- [2] C. L. Cheeseman and S. Harris. Methods of marking badgers (*Meles meles*). *Journal of Zoology*, 197:289–292, 1982.
- [3] C. L. Cheeseman and P. J. Mallinson. *Radio tracking in the study of bovine tuberculosis in badgers*, pages 649–656. Pergamon Press, Oxford and New York, 1979.
- [4] Andrew J. K. Conlan, Trevelyan J. McKinley, Katerina Karolemeas, Ellen Brooks Pollock, Anthony V. Goodchild, Andrew P. Mitchell, Colin P. D. Birch, Richard S. Clifton-Hadley, and James L. N. Wood. Estimating the hidden burden of bovine tuberculosis in Great Britain. *PLOS Computational Biology*, 8(10), 2012.
- [5] Anne Cori, Neil M. Ferguson, Christophe Fraser, and Simon Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology*, 178(9):1505–1512, 2013.
- [6] R. J. Delahay, N. Walker, G. S. Smith, D. Wilkinson, R. S. Clifton-Hadley, C. L. Cheeseman, A. J. Tomlinson, and M. A. Chambers. Long-term temporal trends and estimated transmission rates for *Mycobacterium bovis* infection in an undisturbed high-density badger (*Meles meles*) population. *Epidemiology & Infection*, 141(7):1445–1456, 2013.
- [7] Christophe Fraser. Estimating individual and household reproduction numbers in an emerging epidemic. *PLoS ONE*, 2(8):e758, 2007.
- [8] Radford M. Neal. MCMC using Hamiltonian Dynamics. In *Handbook of Markov Chain Monte Carlo*, chapter 5, pages 113–162. Chapman & Hall/CRC, 2011.
- [9] Gareth O. Roberts and Jeffrey S. Rosenthal. Examples of Adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2):349–367, 2009.
- [10] Matthew J. Smith, Sandra Telfer, Eva R. Kallio, Sarah Burthe, Alex R. Cook, Xavier Lambin, and Michael Begon. Host-pathogen time series data in wildlife support a transmission function between density and frequency dependence. *Proceedings of the National Academy of Sciences*, 106(19):7905–7909, 2009.
- [11] F. A. M. Tuyttens, D. W. Macdonald, R. J. Delahay, L. M. Rogers, P. J. Mallinson, C. A. Donnelly, and C. Newman. Differences in trappability of European badgers *Meles meles* in three populations in England. *Journal of Applied Ecology*, 36:1051–1062, 1996.
- [12] Jacco Wallinga and Peter Teunis. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of Epidemiology*, 160(6):509–516, 2004.