Technical Appendix

In our implementation of the reversible jump MCMC algorithm we update each of the parameters in turn, and then update the augmented data. When updating the vector of β values each element is updated in turn with a random walk proposal. A proposal, $\beta_k^\star$, for a new value for the kth element $\beta_{\mathbf{k}}$ is chosen by drawing from the proposal distribution $N(\beta_{k,c}, \sigma^2_{\beta_k})$. Here $\beta_{k,c}$ is the current value of β_k and $\sigma^2_{\beta_k}$ the proposal variance.

The proposed value, $\beta_k^\star$, is then accepted into the sample with a probability, $\alpha(\beta_{k,c}, \beta_k^\star)$, chosen to ensure the chain has the correct stationary distribution (1):

$$\alpha(\beta_{k,c}, \beta_k^\star) = min\left[1, \frac{p(\mathbf{A}|\beta_{\mathbf{k}}^\star, \nu, \xi, \phi)p(\beta_k^\star)}{p(\mathbf{A}|\beta_{\mathbf{k,c}}, \nu, \xi, \phi)p(\beta_{k,c})}\right].$$

Here $p(\mathbf{A}|\beta, \nu, \xi, \phi)$ is the likelihood of the augmented data for given β, ν, ξ and ϕ; and $p(\beta_k)$ is the prior density for β_k.

The proposal variance, $\sigma^2_{\beta_k}$, acts as a tuning constant: too large and too few moves are accepted; too small, and the new values are too close to the old ones for the Markov chain to mix well.

Updates to ϕ proceed similarly, except that a proposal for $\ln \phi^\star$ is drawn from $N(\ln \phi_c, \sigma^2_\phi)$ (the log transformation ensures a positive value for ϕ, but allows the proposal to be drawn from the whole real line). Because of this change of parameterisation the acceptance ratio becomes:

$$\alpha(\phi_c, \phi^\star) = min\left[1, \frac{p(\mathbf{A}|\beta, \nu, \xi, \phi^\star)p(\phi^\star)\phi^\star}{p(\mathbf{A}|\beta, \nu, \xi, \phi_c)p(\phi_c)\phi_c}\right].$$

where the $\phi^\star$ and ϕ_c terms in the numerator and denominator are the Jacobians for the change of parameterisation.

Updates to ν and ξ are performed by Gibbs sampling (2), except in models VI and VII when a Metropolis step is used to update ν.

Updates to the augmented data Let $\mathbf{A}^\star$ represent the proposed augmented data, and $\mathbf{A_c}$ the current augmented data.

A proposal $\mathbf{A}^\star$ is accepted as the new state of the augmented data with probability

$$\alpha(\mathbf{A_c}, \mathbf{A}^\star) = min(1, H)$$

where

$$H = \left[\frac{p(\mathbf{D}|\mathbf{A}^\star)p(\mathbf{A}^\star|\beta, \nu, \xi, \phi)}{p(\mathbf{D}|\mathbf{A_c})p(\mathbf{A_c}|\beta, \nu, \xi, \phi)}\right]\Pi_p. \quad (1)$$

Π_p is the proposal ratio: the ratio of the density of a proposal for a move from $\mathbf{A}^\star$ to $\mathbf{A_c}$, $q(\mathbf{A}^\star \rightarrow \mathbf{A_c})$, to the density of the proposal for the reverse move, $q(\mathbf{A_c} \rightarrow \mathbf{A}^\star)$.

Four types of changes to the augmented data are used: updates to the times of new acquisitions of the pathogens (M1); adding new importation events for the pathogens (M2); adding new cross-infection events (M3); and switching between importation and cross-infection events (M4). When updating the augmented data one of these moves is chosen at random with equal probability. These updates and associated proposal ratios are described in detail below. It can be shown that these four types of move are sufficient to ensure irreducibility. Aperiodicity follows from irreducibility, since some states will clearly have a periodicity of one.

Updates to the augmented data and associated proposal ratios are described in detail below.

- Move M1: updates to event times.

  In this move only the timing of new acquisitions of the pathogen is changed. This update therefore applies only to patient episodes where patients become colonized in the current augmented data, but when there is uncertainty as to the exact colonization date. This move is carried out on one patient episode at a time, where the episode is randomly chosen from those where colonization is acquired in the current augmented data (i.e. $c_i = 1$). Suppose that when updating the acquisition time for patient episode i, there are n_i possible days on which the strain may have been acquired. Because, once colonized, patients are assumed to remain colonized for the duration of their stay, these n_i days are necessarily contiguous. One of these n_i days is selected uniformly, and in $\mathbf{A}^\star$ patient i is made colonized from the selected day to discharge. This move doesn't change the dimension of the augmented data and hence $\Pi_p = 1$.

- Move M2: addition/removal of importation events.

  This move applies only to patient episodes having no positive swabs, and for episodes where patients are currently assumed in the augmented data to either never carry the organism, or to be carrying the organism when admitted to the ward. Only one patient episode is updated at a time.

  A decision to add an importation is made with probability 0.5. In this case, one of the m_c candidate patients who never carry the organism in

$\mathbf{A_c}$ and who have no swabs taken is chosen at random and in the proposed augmented data, $\mathbf{A}^\star$, this patient is made to be colonized from admission until discharge. Similarly, a decision to remove an importation in the proposed augmented data is made with probability 0.5, and one of the m'_c candidate patients assumed positive on admission in $\mathbf{A_c}$ but with no positive swabs is selected uniformly. The chosen patient is then set to be free of the organism in $\mathbf{A}^\star$.

The proposal ratio, Π_p, for adding an importation is $m_c/(m'_c + 1)$, while for the reverse move it is $m'_c/(m_c + 1)$.

- Move M3: addition/removal of cross-infections.

  This move applies only to patient episodes with no positive swabs, and for episodes where patients are not currently assumed in the augmented data to carry the organism on admission to the ward. Only one patient episode is updated at a time.

  A decision to add or remove a cross-infection is first made (each has a probability of 0.5). If adding a cross-infection, one of the m_c candidate patient episodes (those never positive in $\mathbf{A_c}$ and with no positive swabs) is chosen uniformly. If patient episode i is chosen one of the n_i days during which the patient could feasibly have acquired the organism is then selected at random and in the proposed augmented data, $\mathbf{A}^\star$, this patient is made to be colonized from this day until discharge.

  If removing a cross-infection, a patient episode is selected with equal probability from the m'_c candidates (patient episodes with no positive swabs who become colonized in $\mathbf{A}^\star$, but not from admission). This patient is made to be uncolonized from admission to discharge in $\mathbf{A}^\star$.

  The proposal ratio, Π_p, for adding a cross-infection for patient episode i is $m_c n_i/(m'_c + 1)$, while for removing a cross-infection it is $m'_c/(m_c + 1)n_i$.

- Move M4: switch between importation and cross-infections.

  This move applies only to patient episodes where the organism is assumed to be carried at some point in the current augmented data but where there is uncertainty as to whether or not the patient was positive on admission.

  A decision to replace an importation with a cross-infection is made with probability 0.5 (as is the opposite move). When changing an importation to a cross-infection, a patient episode i is selected with

equal probability from m_c candidate episodes (those assumed to be positive on admission in $\mathbf{A_c}$, but for whom the true admission status is unknown). In the proposed augmented data, $\mathbf{A}^\star$, this patient is then made to be negative on admission, and one of the n_i days during which the patient could feasibly have acquired the organism is chosen at random. The patient is made to be colonized from this day until discharge in the proposed augmented data. The reverse move similarly selects one of the m'_c patient episodes with uncertain admission status currently assumed to be cross-infections, and sets this patient to be colonized on admission for the selected episode in $\mathbf{A}^\star$.

The proposal ratio for replacing an importation with a cross-infection for patient episode i is $m_c n_i/(m'_c + 1)$, while the reverse move has proposal ratio $m'_c/(m_c + 1)n_i$.

References

[1] Gilks W, Richardson S, Spiegelhalter D, eds. Markov Chain Monte Carlo in Practice. London: Chapman and Hall/CRC, 1996.

[2] Gelman A, Carlin J, Stern H, et al. Bayesian Data Analysis. Boca Raton, Florida: Chapman and Hall, 2004, 2nd edition.