
Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/39887/9/Smallpox%20Supplementary%20material2.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the Creative Commons Attribution licence and may be reused according to the conditions of the licence. For more details see:
http://creativecommons.org/licenses/by/2.5/

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk
Modelling and Bayesian Inference of the Abakaliki Smallpox Data: Supplementary Material

We describe (i) for the likelihood, how to integrate out the unknown protection statuses of individuals outside the compound; (ii) the details of the MCMC algorithm. For ease of reference we start by recalling the notation used.

1 Notation

The principal notation used is listed in Table 1, most of which is also used in the main manuscript. We also extend the notation of the main manuscript for sets of individuals; specifically, \( N^a_b \) is used to denote a set of individuals with characteristics \( a \) and \( b \), where \( a \) denotes location (such as \( OC \), meaning outside the compound) and \( b \) denotes status such as FTC or n-FTC (i.e. not FTC), sus (initially susceptible), inf (ever infected) or n-inf (never infected). Note that ‘initially susceptible’ can mean either ‘not vaccinated’ or ‘vaccinated, but not protected’.

2 Transformation of the Likelihood

We now explain how to write the likelihood in a form that does not include the protection statuses for individuals outside of the compounds. In section 2.1 we will show that

\[
\pi(r, \gamma | \Phi) = \pi(r, \tilde{\gamma}, x, y | \Phi),
\]

where instead of separate protection statuses for each outside individual, we only require quantities \( x \) and \( y \) where

\[
\begin{align*}
    x &= \text{Number of FTC, vaccinated, unprotected but never infected individuals outside the compounds} \\
    y &= \text{Number of non-FTC, vaccinated, unprotected but never infected individuals outside the compounds,}
\end{align*}
\]
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>Population size</td>
</tr>
<tr>
<td>$n_{com}$</td>
<td>Number of individuals in the compounds</td>
</tr>
<tr>
<td>$\lambda_a$</td>
<td>Global infection rate</td>
</tr>
<tr>
<td>$\lambda_f$</td>
<td>FTC infection rate</td>
</tr>
<tr>
<td>$\lambda_h$</td>
<td>Compound same-confession infection rate</td>
</tr>
<tr>
<td>$b$</td>
<td>Infectivity factor during fever</td>
</tr>
<tr>
<td>$v$</td>
<td>Vaccine efficacy</td>
</tr>
<tr>
<td>$t_q$</td>
<td>Time quarantine measures introduced</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Fixed parameters for disease progression</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Identity of initial infective</td>
</tr>
<tr>
<td>$e_\kappa$</td>
<td>Exposure time of $\kappa$</td>
</tr>
<tr>
<td>$s$</td>
<td>Vector of vaccination statuses (all individuals)</td>
</tr>
<tr>
<td>$p$</td>
<td>Vector of protection statuses (all individuals)</td>
</tr>
<tr>
<td>$\tilde{p}$</td>
<td>Vector of protection statuses (compound individuals only)</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>$(\kappa, e_\kappa, t_q, b, v, \lambda_a, \lambda_f, \lambda_h, \theta, s)$</td>
</tr>
<tr>
<td>$e$</td>
<td>Vector of exposure times other than $e_\kappa$</td>
</tr>
<tr>
<td>$i$</td>
<td>Vector of fever-period start times</td>
</tr>
<tr>
<td>$r$</td>
<td>Vector of rash times (data)</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Vector of removal times</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Vector of quarantine times</td>
</tr>
<tr>
<td>$\tilde{\gamma}$</td>
<td>$(e, i, q, \tau, p)$</td>
</tr>
<tr>
<td>$\tilde{\gamma}$</td>
<td>$(e, i, q, \tau, \tilde{p})$</td>
</tr>
<tr>
<td>$N_{inf}$</td>
<td>Set of ever-infected individuals</td>
</tr>
<tr>
<td>$N_a^b$</td>
<td>Set of individuals with location $a$ and status $b$</td>
</tr>
</tbody>
</table>

Table 1: Principal notation used in Supplementary Material
and where
\[ \tilde{\gamma} = (e, i, q, \tau, \tilde{p}) \]
where \( \tilde{p} = (p_0, p_1, \ldots, p_{n_{com} - 1}) \). Thus \( \tilde{\gamma} \) is the same as \( \gamma \) but with protection statuses for individuals inside the compounds only.

Given (1), \( x \) and \( y \) can then be integrated out to yield a likelihood which is computationally much faster to calculate. First note that
\[ \pi(r, \tilde{\gamma}, x, y | \Phi) = \pi(r, \tilde{\gamma} | x, y, \Phi) \pi(x, y | \Phi) \]
and hence
\[ \pi(\Phi, \tilde{\gamma}, x, y | r) = \frac{\pi(r, \tilde{\gamma} | x, y, \Phi) \pi(x, y | \Phi) \pi(\Phi)}{\pi(r)}. \]

Integrating out \( x \) and \( y \), which is equivalent to summing in this case since they take discrete values, we obtain the new target density
\[ \pi(\Phi, \tilde{\gamma} | r) = \sum_{x,y} \pi(\Phi, \tilde{\gamma}, x, y | r) = \frac{\pi(\Phi)}{\pi(\Phi)} \sum_{x,y} \pi(r, \tilde{\gamma} | x, y, \Phi) \pi(x, y | \Phi). \]

### 2.1 Derivation of equation (1)

First, note that both \( x \) and \( y \) are determined by the data, the vaccination status of individuals outside the compound, and the protection status of individuals outside the compound. Since only the latter is unknown, we may think of \( x \) and \( y \) as functions of \( p \), or more precisely those elements of \( p \) that correspond to vaccinated individuals outside the compound. In particular, summation over all values of \( x \) and \( y \) is equivalent to summation over such elements of \( p \).

Recall that \( \Lambda(t) \) denotes the total infectious pressure acting on all susceptibles at time \( t \), and that we have
\[
\pi(r, \gamma|\Phi) = \left( \prod_{j \in N_{inf}} \Lambda_j(e_j- \right) \times e^{-\int_{t_0}^T \Lambda(t) dt} \\
\times \prod_{j \in N_{inf}} f_I(i_j - e_j) f_F(r_j - i_j) f_R(\tau_j - r_j) f_Q(q_j - \max(r_j, t_q)) \\
\times \sum_{v=0}^{N-1} p_0 s_v \times \sum_{v=0}^{N-1} (1-p_0) s_v \times (1-v). 
\]

We may write \( \Lambda(t) = \Lambda_{OC}(t) + \Lambda_{CN}(t) + \Lambda_{CC}(t) \), where (i) \( \Lambda_{OC}(t) \) is the pressure on individuals outside the compounds; (ii) \( \Lambda_{CN}(t) \) is the pressure on those inside the compounds who never get infected; and (iii) \( \Lambda_{CC}(t) \) is the pressure on those inside who do get
infected. Then the only terms in (2) with unknown protection statuses for outside individuals are \( \Lambda_{OC}(t) \) and the terms involving \( v \). We now expand the likelihood of the protection status, according to whether each individual is inside/outside and by whether they do/do not become infected, as follows:

\[
\sum_{r'=0}^{N} p_{r'} s_{r'} (1 - v)^{(1 - v)^{r}} \sum_{r=0}^{n_{\text{com}} - 1} p_{r} s_{r} (1 - p_{r}) s_{r} \\
\times \sum_{r'=n_{\text{com}}}^{N} p_{r'} s_{r'} (1 - v)^{r} \sum_{r=n_{\text{com}}}^{n_{\text{com}} - 1} p_{r} s_{r} (1 - p_{r}) s_{r} \\
\times (1 - v)^{r} \sum_{r'=n_{\text{inf}}}^{N} p_{r'} s_{r'} (1 - v)^{r} \sum_{r=n_{\text{inf}}}^{n_{\text{inf}} - 1} p_{r} s_{r} (1 - p_{r}) s_{r} \\
\times (1 - v)^{r} \sum_{r'=n_{\text{inf}} - n_{\text{inf}}}^{N} p_{r'} s_{r'} (1 - v)^{r} \sum_{r=n_{\text{inf}} - n_{\text{inf}}}^{n_{\text{inf}} - n_{\text{inf}} - 1} p_{r} s_{r} (1 - p_{r}) s_{r}.
\]

The terms corresponding to individuals inside the compounds do not depend on \( x \) and \( y \), and neither does the term concerning outside infectives, since their protection status is known. Thus the part of the likelihood dependent on \( x \) and \( y \), which are determined by \( p \) as previously explained, is

\[
L_{x,y}(p) = e^{-\int_{e_{K}}^{T} \Lambda_{OC}(t) dt} \sum_{r'=n_{\text{com}}}^{N} p_{r'} s_{r'} (1 - v)^{r} \sum_{r=n_{\text{com}}}^{n_{\text{com}} - 1} p_{r} s_{r} (1 - p_{r}) s_{r} \\
\times \sum_{r'=n_{\text{com}}}^{N} p_{r'} s_{r'} (1 - v)^{r} \sum_{r=n_{\text{com}}}^{n_{\text{com}} - 1} p_{r} s_{r} (1 - p_{r}) s_{r} \\
\times (1 - v)^{r} \sum_{r'=n_{\text{inf}}}^{N} p_{r'} s_{r'} (1 - v)^{r} \sum_{r=n_{\text{inf}}}^{n_{\text{inf}} - 1} p_{r} s_{r} (1 - p_{r}) s_{r} \\
\times (1 - v)^{r} \sum_{r'=n_{\text{inf}} - n_{\text{inf}}}^{N} p_{r'} s_{r'} (1 - v)^{r} \sum_{r=n_{\text{inf}} - n_{\text{inf}}}^{n_{\text{inf}} - n_{\text{inf}} - 1} p_{r} s_{r} (1 - p_{r}) s_{r}.
\]

The integral of \( \Lambda_{OC} \) is the sum over all infectives \( j \) of the pressure from \( j \) to any given FTC or non-FTC individual outside of the compounds (throughout all time \( (e_{K},T) \)), summed over all the initially susceptible FTC and non-FTC individuals. Thus,

\[
e^{-\int_{e_{K}}^{T} \Lambda_{OC}(t) dt} = \exp \left( - \sum_{j \in N_{\text{inf}}}^{N_{\text{inf}}} \sum_{k \in N_{\text{inf}}^\text{FTC}} \Psi_{jk} \right)
\]

with \( \Psi_{jk} = \text{total infectious pressure from } j \text{ to susceptible } k \text{ during the time interval } (e_{K},T) \).

We now expand the right-hand side of (3) according to whether susceptibles become infected or not, and whether they are FTC or not, yielding

\[
\exp \left( - \sum_{j \in N_{\text{inf}}}^{N_{\text{inf}}} \sum_{k \in N_{\text{inf}}^\text{FTC}} \Psi_{jk} \right) = \exp \left( - \sum_{j \in N_{\text{inf}}}^{N_{\text{inf}}} \left( \sum_{k \in N_{\text{inf}}^\text{FTC}} \Psi_{jk} + \sum_{k \in N_{\text{inf}}^\text{n-FTC}} \Psi_{jk} \right) \right)
\]

\[
= \exp \left( - \sum_{j \in N_{\text{inf}}}^{N_{\text{inf}}} \left( \sum_{k \in N_{\text{inf}}^\text{FTC}} \Psi_{jk} + \sum_{k \in N_{\text{inf}}^\text{n-FTC}} \Psi_{jk} \right) \right)
\]

\[
+ \sum_{k \in N_{\text{inf}}^\text{n-FTC}} \Psi_{jk} \right)
\]
Recall that the data only tell us if individuals are vaccinated, not whether they are protected. Of the four $\Psi_{jk}$ values in equation (5), the numbers of initially susceptible FTC and non-FTC individuals outside that do become infective are known, but not the number of initial susceptibles that are never infected. Figure 1 displays how individuals can be in the latter two categories.

So, the total number of initially susceptible, never infected individuals outside the compounds is given by

$$|N^{oc}_{n-inf}| = a^{oc}_{n-inf, FTC} + a^{oc}_{n-inf, n-FTC} + x + y$$

where $a^{oc}_{n-inf, FTC}$ is the known number of non-vaccinated, never infective FTC individuals outside the compounds and $a^{oc}_{n-inf, n-FTC}$ is the corresponding number of non-FTC individuals. Using this and (5), (3) becomes

$$L_{x,y}(\mathbf{p}) = \exp \left( - \sum_{j \in N_{inf}} \left( \sum_{k \in N^{oc}_{inf, FTC}} \Psi_{jk} + \sum_{k \in N^{oc}_{inf, n-FTC}} \Psi_{jk} \right) + \chi_F(j) \times (a^{oc}_{n-inf, FTC} + x) + \chi_N F(j) \times (a^{oc}_{n-inf, n-FTC} + y) \right)$$

$$\times \prod_{r \in N_{com}} p_r s_r \times \prod_{r \in N_{n-inf}} (1 - p_r) s_r$$

$$\times \prod_{r \in N_{n-inf}} (1 - v)^{N_{n-inf}}$$

(6)
where

\[ \chi_F(j) = (b(r_j - i_j) + (\min(q_j, \tau_j) - r_j)) \times \begin{cases} \frac{\lambda_a}{N-1} + \frac{\lambda_f}{N-1} & \text{if } f_j = \text{FTC} \\ \frac{\lambda_a}{N-1} & \text{otherwise} \end{cases} \]

and

\[ \chi_{NF}(j) = (b(r_j - i_j) + (\min(q_j, \tau_j) - r_j)) \times \begin{cases} \frac{\lambda_a}{N-1} & \text{if } f_j = \text{FTC} \\ \frac{\lambda_a}{N-1} + \frac{\lambda_f}{N-1} & \text{otherwise} \end{cases} \]

represent respectively the contribution from infective \( j \) to a never infected FTC and non-FTC susceptible outside the compounds over all time \((e_k, T)\), this contribution being equal for all susceptibles \( k \).

Considering the protection status likelihood parts of equation (6), note that

\[ (1 - v)^N \sum_{r=0}^{N} (1-p_r)s_r = (1 - v)^x+y, \]

since the sum is equal to the number of vaccinated but unprotected individuals outside who do not become infected, and hence

\[ \sum_{r=0}^{N} \left( \sum_{r=0}^{N} p_r \right) s_r = b^{\text{oc}}_n - inf, \text{FTC}, \]

where \( b^{\text{oc}}_n - inf, \text{FTC} \) is the number of never-infected vaccinated individuals outside the compounds.

Splitting up the terms involving \( x \) and \( y \), equation (6) can be written in the form

\[ L_{x, y}(\mathbf{p}) = \exp \left( - \sum_{j \in N_{inf}} \left( \sum_{k \in N_{inf}, FTC} \Psi_{jk} + \sum_{k \in N_{inf}, n \neq FTC} \Psi_{jk} + a_{n - inf, FTC}^{\text{oc}} \chi_F(j) + a_{n - inf, n \neq FTC}^{\text{oc}} \chi_{NF}(j) \right) \right) \times \exp \left( - \sum_{j \in N_{inf}} \chi_F(j) \times x \right) \times (1 - v)^x b^{\text{oc}}_n - inf, FTC - x \times \exp \left( - \sum_{j \in N_{inf}} \chi_{NF}(j) \times y \right) \times (1 - v)^y b^{\text{oc}}_n - inf, n \neq FTC - y \]

which establishes (1) as claimed, the point being that none of the unknown protection statuses for outside individuals are explicitly required.
2.2 Removing $x$ and $y$

We now sum (7) over $x$ and $y$. Recall that this is equivalent to summing $L_{x,y}(p)$ over all possible values of the elements of $p$ relating to vaccinated individuals outside the compound. For fixed $x$ and $y$, write $p(x,y)$ for a typical vector $p$ that yields $x$ and $y$. First note that the number of possible vectors $p(x,y)$ is

$$\binom{b_{oc}}{x} \binom{b_{oc}}{y},$$

where $b_{oc}^{n-inf, FTC}$ and $b_{oc}^{n-inf,n-FTC}$ are respectively the (known) number of vaccinated, not-infected FTC and non FTC individuals outside the compounds. We may thus write

$$L_{x,y} = \sum_{p(x,y)} L_{x,y}(p) = \binom{b_{oc}}{x} \binom{b_{oc}}{y} L_{x,y}(p),$$

which in combination with (7) means that the required sums (over $x$ and $y$, which can be separated from each other) are of the form

$$\sum_{x=0}^{b} \binom{b}{x} e^{-\alpha x} (1-v)^x v^{b-x} = ((1-v)e^{-\alpha} + v)^b.$$

Writing $L = \sum_{x,y} L_{x,y}$, $\chi_F = \sum_{j \in N_{inf}} \chi_F(j)$ and $\chi_{NF} = \sum_{j \in N_{inf}} \chi_{NF}(j)$ we thus obtain

$$L = \exp \left( - \sum_{j \in N_{inf}} \left( \sum_{k \in N_{inf, FTC}} \Psi_{jk} + \sum_{k \in N_{inf,n-FTC}} \Psi_{jk} \right) - a_{oc}^{n-inf, FTC} \chi_F - a_{oc}^{n-inf,n-FTC} \chi_{NF} \right) \times \left( (1-v)e^{-\chi_F} + v \right)^{b_{oc}^{n-inf, FTC}} \left( (1-v)e^{-\chi_{NF}} + v \right)^{b_{oc}^{n-inf,n-FTC}}.$$

To obtain the full likelihood expression, we combine this with the remaining parts of the original likelihood from equation (2). Finally, for numerical stability in the MCMC algo-
Algorithm we work with log-likelihood, so taking logs gives an overall log likelihood of

\[
\log(\pi(r, \tilde{\gamma} | \Phi)) = \log \left( \prod_{j \in N_{inf}} \Lambda_j(e_j^-) \right) - \int_{t_{\kappa}}^T (\Lambda_{CN}(t) - \Lambda_{CC}(t)) \, dt - \sum_{j \in N_{inf}} \left( \sum_{k \in N_{inf}^{oc}} \Psi_{jk} \right) - a_{n-inf, FTC}^{oc} \chi_{F} - a_{n-inf,n-FTC}^{oc} \chi_{NF} + b_{n-inf, FTC}^{oc} \log \left( v + (1 - v)e^{-\chi_{F}} \right) + b_{n-inf,n-FTC}^{oc} \log \left( v + (1 - v)e^{-\chi_{NF}} \right) + \log \left( \prod_{j \in N_{inf}} f_I(i_j - e_j)f_F(r_j - i_j)f_R(\tau_j - r_j)f_Q(q_j - \max(r_j, t_q)) \right) + \sum_{r=0}^{n_{com}-1} s_r(p_r \log(v) + (1 - p_r) \log(1 - v)) + \sum_{r=n_{com}}^{N-1} s_r \log(1 - v).
\]

3 MCMC

Here we give details of the MCMC algorithm.

3.1 Unknown vaccination statuses in MCMC algorithm

Although the total number of vaccinated individuals within the Abakaliki compounds is known, we do not have complete information on the composition of compound individuals with respect to vaccination status and FTC membership. Table 2 shows 5 potential configurations of the 12 individuals with unknown details. This means that all the elements of \( s \) (the vector of vaccination statuses) are assumed known, with 12 exceptions, which can take 5 possible configurations. A described below in the MCMC algorithm, updates to \( s \) involve potentially changing the configuration from the 5 that are possible, while all other elements of \( s \) are unchanged.

3.2 Full conditional Distributions

Here we give, up to proportionality, the full conditional distributions for the parameters that will be updated in the MCMC algorithm. Although in principle we could update all parameters individually using a Metropolis-Hastings step involving the full likelihood, algorithm efficiency is increased if we only include those parts of the likelihood that can change during a particular update.

The parameters to be updated are \( t_q, \lambda_a, \lambda_f, \lambda_h, v, b, \tilde{p}, s \) and the exposure, fever, removal and quarantine times. In the following, \( L \) is defined in equation 8, and \( c \) denotes a
Table 2: Possible combinations of twelve individuals, labelled 183, 213, 214, 215, 217, 231, 232, 233, 234, 236, 237, 250, with unknown vaccination status

<table>
<thead>
<tr>
<th>Combination</th>
<th>Compound</th>
<th>FTC</th>
<th>Non-FTC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Vaccinated</td>
<td>Nonvaccinated</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>213</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>215, 214</td>
<td>232, 233, 234</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>214</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>213</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>215</td>
<td>214</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>232, 233, 234</td>
<td>236, 237, 250</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>213</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>215</td>
<td>214</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>234, 232, 233</td>
<td>236, 237, 250</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>213</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>214, 215</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>234, 232, 233</td>
<td>236, 237, 250</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>213</td>
<td>183</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>214, 215</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>234, 233</td>
<td>232</td>
</tr>
</tbody>
</table>
generic unknown constant (arising from the fact that we generally only know full conditional distributions up to proportionality).

\[
\log(\pi(t_q | r, \theta, \tilde{\gamma}, \kappa, b, v, \lambda_a, \lambda_f, \lambda_h, e_\kappa, s)) + c = \\
\log \left( \prod_{j \in N_{in, f}} \Lambda_j(e_j-) \right) - \int_{\epsilon_\kappa}^T (\Lambda_{CN}(t) + \Lambda_{CC}(t)) \, dt - \log(L) + \log(\pi(t_q)).
\]

\[
\log(\pi(b | r, \theta, \tilde{\gamma}, \kappa, t_q, v, \lambda_a, \lambda_f, \lambda_h, e_\kappa, s)) + c = \\
\log \left( \prod_{j \in N_{in, f}} \Lambda_j(e_j-) \right) - \int_{\epsilon_\kappa}^T (\Lambda_{CN}(t) + \Lambda_{CC}(t)) \, dt - \log(L) + \log(\pi(b)).
\]

\[
\log(\pi(v | r, \theta, \tilde{\gamma}, \kappa, t_q, b, \lambda_a, \lambda_f, \lambda_h, e_\kappa, s)) + c = \\
- \log(L) + \sum_{r=0}^{n_{com}-1} p_r s_r \log(v) + \sum_{r=0}^{n_{com}-1} (1-p_r) s_r \log(1-v) \\
+ \sum_{r=n_{com}}^{N-1} (1-p_r) s_r \log(1-v) + \log(\pi(v)).
\]

\[
\log(\pi(\lambda_a | r, \theta, \tilde{\gamma}, \kappa, t_q, v, b, \lambda_f, \lambda_h, e_\kappa, s)) + c = \\
\log \left( \prod_{j \in N_{in, f}} \Lambda_j(e_j-) \right) - \int_{\epsilon_\kappa}^T (\Lambda_{CN}(t) + \Lambda_{CC}(t)) \, dt - \log(L) + \log(\pi(\lambda_a)).
\]

\[
\log(\pi(\lambda_f | r, \theta, \tilde{\gamma}, \kappa, t_q, v, b, \lambda_a, \lambda_h, e_\kappa, s)) + c = \\
\log \left( \prod_{j \in N_{in, f}} \Lambda_j(e_j-) \right) - \int_{\epsilon_\kappa}^T (\Lambda_{CN}(t) + \Lambda_{CC}(t)) \, dt - \log(L) + \log(\pi(\lambda_f)).
\]

\[
\log(\pi(\lambda_h | r, \theta, \tilde{\gamma}, \kappa, t_q, v, b, \lambda_f, \lambda_a, e_\kappa, s)) + c = \\
\log \left( \prod_{j \in N_{in, f}} \Lambda_j(e_j-) \right) - \int_{\epsilon_\kappa}^T (\Lambda_{CN}(t) + \Lambda_{CC}(t)) \, dt + \log(\pi(\lambda_h)).
\]
For any $i = 0, 1, \ldots, n_{com} - 1$, defining $\tilde{p}_i^* = (p_0, p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_{n_{com}-1})$,

$$\log(\pi(\tilde{p}_i | r, \Phi, e, i, q, \tau, \tilde{p}_i^*)) + c = -\int_{e_\kappa}^{T} \Lambda_{CN}(t) \, dt + p_i \log(v) + (1 - p_i) \log(1 - v).$$

$$\log(\pi(s | r, \theta, \tilde{\gamma}, \kappa, t_q, v, b, \lambda_a, \lambda_h, e_\kappa)) + c = -\int_{e_\kappa}^{T} \Lambda_{CN}(t) \, dt + \sum_{r=0}^{n_{com}-1} p_r s_r \log(v) + \sum_{r=0}^{n_{com}-1} (1 - p_r) s_r \log(1 - v) + \sum_{r=n_{com} - inf}^{N-1} (1 - p_r) s_r \log(1 - v).$$

Finally, for $i = 0, 1, \ldots, N-1$ where $i$ is infective, defining $\tilde{e}_i^* = (e_0, e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_{N-1})$ and similarly defining $\tilde{i}_i^*, \tilde{q}_i^*$ and $\tilde{\tau}_i^*$,

$$\log(\pi(e_i | r, \Phi, e_i^*, i, q, \tau, \tilde{p})) + c = \log\left(\prod_{j \in N_{inf}} \Lambda_j(e_j,-)\right) - \int_{e_\kappa}^{T} (\Lambda_{CN}(t) + \Lambda_{CC}(t)) \, dt - \log(L).$$

$$\log(\pi(i_i | r, \Phi, e_i^*, i, q_i^*, \tau, \tilde{p})) + c = \log\left(\prod_{j \in N_{inf}} \Lambda_j(e_j,-)\right) - \int_{e_\kappa}^{T} (\Lambda_{CN}(t) + \Lambda_{CC}(t)) \, dt - \log(L).$$

$$\log(\pi(q_i | r, \Phi, e, i, q_i^*, \tau, \tilde{p})) + c = \log\left(\prod_{j \in N_{inf}} \Lambda_j(e_j,-)\right) - \int_{e_\kappa}^{T} (\Lambda_{CN}(t) + \Lambda_{CC}(t)) \, dt - \log(L).$$

$$\log(\pi(\tau_i | r, \Phi, e, i, q_i^*, \tau_i^*, \tilde{p})) + c = \log\left(\prod_{j \in N_{inf}} \Lambda_j(e_j,-)\right) - \int_{e_\kappa}^{T} (\Lambda_{CN}(t) + \Lambda_{CC}(t)) \, dt - \log(L).$$
3.3 MCMC Algorithm

The MCMC algorithm consists of individual Metropolis-Hastings updates for all of the required parameters. For a generic parameter $\alpha$ the update consists of a proposed new value $\alpha^*$ drawn from proposal density (or mass function) $q(\alpha^*|\alpha)$, which is then accepted with probability

$$\min \left( \frac{\pi(\alpha^*) q(\alpha|\alpha^*)}{\pi(\alpha) q(\alpha^*|\alpha)}, 1 \right),$$

where $\pi(\alpha^*)$ and $\pi(\alpha)$ denote the full conditional density (or mass function) for $\alpha$ evaluated at the proposed and current value, respectively. In most cases below, the proposal density $q$ is symmetric in the sense that $q(\alpha|\alpha^*) = q(\alpha^*|\alpha)$, which simplifies the computations.

For $t_q$, $\lambda_a$, $\lambda_f$, $\lambda_h$, $v$, and $b$, we use a Gaussian proposal distribution with pre-specified variance, centered on the current parameter value. Proposed values outside the possible set for each parameter are immediately rejected (for example, all the infection rates must be positive, $0 \leq v \leq 1$, etc.) Note that if such a rejection occurs, no further proposal attempt is made. The acceptance probability is calculated using the full conditional distributions listed above.

Protection status updates (i.e. updates to $\tilde{p}$) are achieved by first randomly selecting a fixed number (e.g. 5 or 10) of compound individuals among those who were vaccinated but never infected, and proposing to change the protection status of each selected individual.

Updates to $s$ are similar: a proposed new value is chosen uniformly at random from one of the five possible configurations of the 12 individuals whose vaccination status is uncertain, as shown in Table 2. Following this, all of the individuals among the 12 who are now proposed to be vaccinated are, independently with probability $v$, proposed to be protected, while those proposed to be unvaccinated are proposed to be unprotected.

Finally, updates to the event times are carried out in pairs, the reason being that we expect these quantities to be correlated. Specifically, an individual is chosen uniformly at random from the set of infected individuals, and proposed new exposure and fever times are simulated from the distributions assumed in the model. These are then accepted or rejected as usual. Following this, a pair of removal and quarantine times are proposed for the same individual, regardless of whether or not the exposure and fever times were changed. In practice we repeat the entire event-time-updating procedure a number of times (e.g. 5) within each full MCMC iteration in order to improve the overall algorithm mixing.

The results in the paper are based on output of 100,000 samples, taken after an initial burn-in of 10,000 iterations and thinning to record every 10th iteration. The algorithm was written in C and took around 2-3 hours to run.