1. Web Appendix A

All of the models which we consider are continuous time models and we use continuous time Markov chain theory. If \( P_h \) is a matrix of probabilities constructed from the relevant expressions given by (2), (3) or (5) the infinitesimal generator, \( G \), is given by

\[
G = \lim_{h \to 0} \frac{1}{h} (P_h - I),
\]

Grimmett and Stirzaker (2001, pp. 258). Firstly we describe how to calculate the likelihood for the bivariate Markov process used to model the prevalence data. Assuming that the first observation is recorded at time \( t \), the joint probabilities at this time can be computed using the matrix exponential

\[
\delta_B e^{tG_B} = P((Y_p(t), Y_h(t) = 0), \ldots, (Y_p(t), Y_h(t) = N)|\delta_B),
\]

where \((Y_p(t), Y_h(t) = j)\) denotes the joint events \((Y_p(t) = 0, Y_h(t) = j), \ldots, (Y_p(t) = N, Y_h(t) = j)\), \( G_B \) is the generator matrix of the prevalence model and \( \delta_B \) denotes the stationary distribution. The likelihood of the observation, the prevalence of colonised patients at time \( t \), \( y_p(t) \), can be computed by summing over all possible values of the healthcare workers prevalence, \( Y_h(t) \), in the above joint probabilities:

\[
P(Y_p(t) = y_p(t)|\delta_B) = \sum_{j=0}^{N} P(Y_p(t) = y_p(t), Y_h(t) = j|\delta_B).
\]

To proceed, a new initial distribution has to be calculated based on the observation, \( y_p(t) \), being known. This can be computed easily using Bayes law of probability which involves the likelihood already calculated. Denoting the new initial distribution given the observation \( Y_p(t) = y_p(t) \) as \( \delta_B(y_p(t)) = P(Y_h(t)|Y_p(t) = y_p(t), \delta_B) \), we find that the elements of \( \delta_B(y_p(t)) \) are

\[
\delta_{Bj}(y_p(t)) = \frac{P(Y_p(t) = y_p(t), Y_h(t) = j|\delta_B)}{P(Y_p(t) = y_p(t)|\delta_B)}, \quad j \in 0, 1, \ldots, N.
\]
The updated initial distribution for the subsequent time interval becomes \((0, \ldots, 0, \delta_B(y_p(t)), 0, \ldots, 0)\) where \(\delta_B(y_p(t))\) is in the appropriate place that has \(Y_p(t) = y_p(t)\).

2. Web Appendix B

For the possibility of allowing imperfect swabbing detection of MRSA, we introduce a Binomial hidden Markov model. The theory is an extension of Cooper and Lipsitch (2004) to the bivariate process, (2) and binomial observation model. The following can be used to calculate the likelihood

\[
L = \delta_B \pi_1(y_1) \Gamma_{t_1} \pi_2(y_2) \Gamma_{t_2 - t_1} \ldots \Gamma_{t_n - t_{n-1}} \pi_n(y_n) 1',
\]

where \(\Gamma_t\) is the transition probability matrix that can be computed using the matrix exponential \(e^{tG_B}\). Since for this dataset the time interval between observations is constant, the matrix exponential only has to be computed once. Here \(\pi_t(y_t)\) is the observation matrix for the \(t^{th}\) observation and is given by

\[
\pi_t(s) = \text{diag}(P(Y(t) = s|Y_p(t) = 0, Y_h(t) = 0), \ldots, P(Y(t) = s|Y_p(t) = 0, Y_h(t) = N), P(Y(t) = s|Y_p(t) = 1, Y_h(t) = 0), \ldots, P(Y(t) = s|Y_p(t) = N, Y_h(t) = N)),
\]

where, in this case, the Binomial probabilities are independent of the prevalence count of colonised health-care workers, \(Y_h(t)\), so that \(P(Y(t) = s|Y_p(t) = i, Y_h(t) = j) = P(Y(t) = s|Y_p(t) = i)\). In order to calculate the Binomial observation probabilities, we use the value of \(Y_p(t)\) for the number of trials after the current time interval and \(\theta_s\) is the probability of success (detecting a colonised patient). Note that to compute the observation probability matrix \(\pi_t(y_t)\) we need to consider all Binomial probabilities of the form \(P(Y(t) = y_t|Y_p(t) = i)\) for \(i \in 0, \ldots, N\). Therefore we need to compute probabilities with the number of trials equal to zero and the number of successes greater than the number of trials. These probabilities are
given by

\[ P(Y(t) = 0|Y_p(t) = 0) = 1, \]
\[ P(Y(t) > i|Y_p(t) = i) = 0, \text{ for } i \in 0, \ldots, N - 1. \]

3. Web Appendix C

The full trivariate process, (3), contains an absorbing state, \( N(t) = M \), however the bivariate chain involving \( Y_p(t) \) and \( Y_h(t) \) remains irreducible. By denoting \( (Y_p(t), Y_h(t), N(t) = n_t) \) to be equivalent to the joint events \( ((Y_p(t) = 0, Y_h(t) = 0, N(t) = n_t), (Y_p(t) = 1, Y_h(t) = 0, N(t) = n_t), \ldots, (Y_p(t) = N, Y_h(t) = 0, N(t) = n_t), (Y_p(t) = 0, Y_h(t) = 1, N(t) = n_t), \ldots, (Y_p(t) = N, Y_h(t) = N, N(t) = n_t)) \), the initial distribution of the trivariate chain is given by

\[ P((Y_p(0), Y_h(0), N(0) = 0); \ldots; (Y_p(0), Y_h(0), N(0) = M)) = (\delta_B, 0, \ldots, 0). \]

To obtain the probabilities at time \( t \) the initial distribution is multiplied by the matrix exponential of the trivariate generator, \( G_C \), multiplied by the time interval \( e^{tG_C} = P((Y_p(t), Y_h(t), N(t) = 0); \ldots; (Y_p(t), Y_h(t), N(t) = M)|N(0) = 0). \]

To compute the likelihood for the observation, the number of new colonisation cases observed in \((0, t)\), denoted by \( n_t \), we marginalise over all possible values of \( Y_p \) and \( Y_h \):

\[ P(N(t) = n_t|N(0) = 0) = \sum_{i=0}^{N} \sum_{j=0}^{N} P(Y_p(t) = i, Y_h(t) = j, N(t) = n_t|N(0) = 0). \]

The new initial distribution has to be found given that we have observed \( N(t) = n_t. \) The new initial distribution, \( \delta_B(n_t) = P(Y_p(t), Y_h(t)|N(t) = n_t, N(0) = 0), \) is easily calculated using Bayes law. Its elements are given by

\[ \delta_B(ij)(n_t) = \frac{P(Y_p(t) = i, Y_h(t) = j, N(t) = n_t|N(0) = 0)}{P(N(t) = n_t|N(0) = 0)}, \quad i, j = 1, \ldots, N. \]

To incorporate the fact that \( N(t) \) is reset to zero at the beginning of each time interval, the new initial distribution for the next time interval is \( (\delta_B(n_t), 0, \ldots, 0) \). This process is
repeated for each observation and the loglikelihood of the data is calculated as the sum of the loglikelihood of each observation. To calculate the likelihood for the pseudo-equilibrium bivariate approximation, (5), to the trivariate process, the same principles as above (ignoring $Y_h(t)$) need to be applied.

4. Web Appendix D

The analysis in the paper assumed that the number of patients in the ICU remained constant (i.e. a discharged patient is immediately replaced by another). Realistically, there are fluctuations in ward utilisation, which the models of the paper did not account for. In this appendix we outline methodology which extends that of the paper to allow the number of patients in the ward to change over time.

Allowing the number of patients in the ward to be a function of time, $N_p(t)$, and assuming a constant health-care worker population, $N_h$, the bivariate model of colonised patients and health-care workers becomes

$$
P \left( \begin{array}{c} Y_p(t+a) = i + 1, \\ Y_h(t+a) = j \end{array} \right| \begin{array}{c} Y_p(t) = i, \\ Y_h(t) = j \end{array} \right) = \phi_1(N_p(t) - i)ja + o(a),
$$

$$
P \left( \begin{array}{c} Y_p(t+a) = i, \\ Y_h(t+a) = j + 1 \end{array} \right| \begin{array}{c} Y_p(t) = i, \\ Y_h(t) = j \end{array} \right) = \phi_2(N_h - j)ia + o(a),
$$

$$
P \left( \begin{array}{c} Y_p(t+a) = i, \\ Y_h(t+a) = j - 1 \end{array} \right| \begin{array}{c} Y_p(t) = i, \\ Y_h(t) = j \end{array} \right) = \kappa ja + o(a).
$$

In the above model, terms involving patients leaving and arriving are not required since the exact discharge and admission times for specific patients are used. Prevalences therefore need to be recorded when the ward size changes.

There are several disadvantages with this approach. The likelihood of the data now involves
dealing with a Markov chain with a state space changing over time and methodology needs to be developed to account for this. In terms of computational burden, the matrix exponential needs to be recalculated every time the number of patients changes. The computational complexity of this model makes it infeasible for the trivariate process that models incidence. 

There are additional disadvantages in using this approach in terms of hospital resources. The approach requires patient specific data and then a significant amount of patient bookkeeping is required to maintain exact arrival and discharge times, rather than just counting incidence over a pre-specified time period.

References
