Transmissibility of Livestock-associated Methicillin-Resistant *Staphylococcus aureus*

Technical Appendix

Statistical Methods

The model of Bootsma et al. (1) predicts that the detected size of an outbreak is geometrically distributed. We denote the parameter of the geometric distribution by \( \xi \). This means that the likelihood that a randomly chosen detected outbreak has size \( i \) (including the index case) is given by \( \xi(1-\xi)^{i-1} \).

If there are \( N \) outbreaks with a total of \( M \) secondary cases, the likelihood of observing these outbreak sizes is given by \( L = \xi^N(1-\xi)^M \). The maximum likelihood estimator, i.e., the value of the parameter \( \xi \) which makes the observations most likely, is given by \( \xi_{MLE} = N/(N + M) \).

A confidence interval for \( \xi_{MLE} \) can be obtained by the profile likelihood method. This confidence interval contains all values for the parameter \( \xi \) of the geometric distribution for which the observations are still sufficiently likely. The cutoff values to determine whether the data are still sufficiently likely depends on whether we are calculating 90% confidence intervals, 95% or 99% confidence intervals and is based on the chi-square distribution.

When we know the discharge rate and the rate at which colonization is detected, we can calculate a ratio of these two. With this ratio \( (r) \) we can translate values for \( \xi \) into values of \( R_A \), the per admission reproduction number by using the formula \( R_A = (1-\xi)(r + \xi) / \xi \) (see Bootsma et al. for more details).

To check whether the model assumptions, which lead to a geometric distribution of the outbreak sizes, are in agreement with the data, we tested whether the observed outbreak sizes are indeed similar to a geometric distribution by performing the Anderson Darling goodness of fit
test. This test is based on the test statistic

\[ A = \frac{1}{\sum_{i=1}^{\infty} p(i) (\hat{F}(i) - F(i))^2} \]

where \( p(i) = \tilde{\xi}_{MLE}(1 - \tilde{\xi}_{MLE})^{i-1} \) is the probability density function of the geometric distribution, \( F(i) \) is the cumulative density function corresponding to \( p(i) \) and \( \hat{F}(i) \) is the empirical cumulative density function.

Reference


Technical Appendix Table A. Mathematical parameters of index cases*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>LA-MRSA, n = 40</th>
<th>Other (non LA)-MRSA, n = 101</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discharge rate (days)</td>
<td>1/13</td>
<td>1/10</td>
</tr>
<tr>
<td>Detection rate (days)</td>
<td>1/20</td>
<td>1/20</td>
</tr>
<tr>
<td>( \xi ) (95% CI)</td>
<td>0.93 (0.83–0.98)</td>
<td>0.70 (0.62–0.77)</td>
</tr>
<tr>
<td>( R_a ) (95% CI)</td>
<td>0.12 (0.03–0.30)</td>
<td>0.52 (0.38–0.69)</td>
</tr>
</tbody>
</table>

*LA-MRSA, livestock-associated meticillin-resistant Staphylococcus aureus.

Technical Appendix Figure. spa-based minimal spanning tree, including spa types of index cases and spa types considered livestock associated. spa types with the same color are considered to be related.