Using a Longitudinal Model to Estimate the Effect of Methicillin-resistant
Staphylococcus aureus Infection on Length of Stay in an Intensive Care Unit

Adrian G. Barnett, Rahul Batra, Nicholas Graves, Jonathan Edgeworth, Julie Robotham, and Ben Cooper

Initially submitted January 18, 2009; accepted for publication July 20, 2009.

Health-care-associated methicillin-resistant Staphylococcus aureus (MRSA) infection may cause increased hospital stay, or sometimes death. Quantifying this effect is complicated because the exposure is time dependent: infection may prolong hospital stay, while longer stays increase infection risk. In this paper, the authors overcome these problems by using a multinomial longitudinal model to estimate the daily probability of death and discharge. They then extend the basic model to estimate how the effect of MRSA infection varies over time and to quantify number of excess days in the intensive care unit due to infection. They found that infection decreased the relative risk of discharge (relative risk ratio = 0.68, 95% credible interval: 0.54, 0.82). Infection on the first day of admission resulted in a mean extra stay of 0.3 days (95% credible interval: 0.1, 0.5) for a patient with an Acute Physiology and Chronic Health Evaluation II score of 10 and 1.2 days (95% credible interval: 0.5, 2.0) for a patient with a score of 30. The decrease in the relative risk of discharge remained fairly constant with day of MRSA infection but was slightly stronger closer to the start of infection. Results confirm the importance of MRSA infection in increasing stay in an intensive care unit but suggest that previous work may have systematically overestimated the effect size.

Estimating additional length of hospital stay due to nosocomial infections creates a number of statistical challenges (5). The central difficulty arises from the fact that infections may increase the length of stay, and increased length of stay simultaneously increases the chance of infection (6). However, most standard regression analyses assume a one-way direction of causation from exposure (infection) to response (length of stay). Standard survival analysis of hospital stay data is also inappropriate because censoring of hospital stays due to death does not occur at random. Instead, the most severely ill will have the highest chance of dying and the lowest chance of being discharged on a given day. Such informative censoring violates the assumptions of standard survival analyses and can lead to very large biases if unaccounted for. A further problem is that factors that may predispose a patient to infection (such as use of invasive

Abbreviations: APACHE, Acute Physiology and Chronic Health Evaluation; CI, credible interval; ICU, intensive care unit; MRSA, methicillin-resistant Staphylococcus aureus; RRR, relative risk ratio; TISS, therapeutic intervention scoring system.
interventions) may also be independently associated with increased length of stay.

In this paper, we aim to address all these challenges. We use individual patient-level data to quantify the additional length of stay in an intensive care unit (ICU) because of infections with one of the most virulent and widely distributed nosocomial pathogens, methicillin-resistant *Staphylococcus aureus* (MRSA).

We make use of a multinomial regression model with a longitudinal data format. This format has the advantage of allowing use of time-dependent exposures (i.e., exposures such as assisted ventilation that may change during a patient’s stay). Time-dependent exposures can be more informative than time-independent exposures because they offer the chance to examine the order of exposure and response.

A further advantage of the longitudinal format is that the effect of time-dependent exposures can be allowed to change over time by using time-dependent covariates. If modeled with only time-independent covariates, a hospital-acquired infection would be assumed to decrease the risk of discharge by the same amount regardless of the length of time the patient had already spent in the ward. In contrast, time-dependent covariates allow changes in the exposure effects over time. Thus, for example, the effect of a nosocomial infection on the chances of being discharged might be strongest on the actual day of infection and then wane over time.

The use of time-dependent exposures has been studied extensively in the field of longitudinal analyses (7, chapter 12). In this paper, we apply some of this knowledge to length-of-stay data.

**MATERIALS AND METHODS**

**Clinical setting**

Data were collected from all patients admitted to 2 adjacent 15-bed ICUs in a 1,100-bed, dual-site teaching hospital in the United Kingdom between January 1, 2002, and April 20, 2006. During the study period, 60% of the patients were medical, 21% surgical, and 19% cardiothoracic. Data comprised age, sex, date of admission and discharge to the ICU, specialty, day 1 Acute Physiology and Chronic Health Evaluation (APACHE)-II score, daily therapeutic intervention scoring system (TISS) score (8) that included measurements required to diagnose a systemic inflammatory response syndrome, dates of starting or stopping ventilation or hemofiltration, date of collection and culture of MRSA from all microbiologic samples, and date of starting treatment with vancomycin or linezolid. Clinical samples were taken only when local or systemic infection was suspected. Further details on infection control and laboratory practice have been previously reported (9, 10).

A MRSA infection was considered to be present only if 3 conditions were satisfied: 1) MRSA was isolated from a sterile or nonsterile clinical sample including a removed vascular catheter tip; 2) there was treatment with vancomycin or linezolid, the only antibiotics used for initial treatment of MRSA; and 3) there was a systemic inflammatory response syndrome response, requiring that 2 of the following criteria be present between 2 days before and 3 days after the positive culture: temperature <36°C or >38°C; heart rate >90 beats/minute; respiratory rate >20 breaths/minute or partial pressure of carbon dioxide <32 mm Hg; and white blood cell count >12,000 cells/mm³ or <4,000 cells/mm³ (11). Patients from whom MRSA was isolated from any site but who did not fulfill these additional criteria were considered to be colonized with MRSA.

**Data quality control**

Data quality control mechanisms included automated range, logic, and date checks. The integrity of the data extraction process was validated for completeness and accuracy by manually comparing 5% of the electronic database with the original source data.

**Exposures**

We were motivated to fit a longitudinal model because the data have a number of important time-dependent exposures. The most important was the presence of a MRSA infection because the primary research question was the impact of such infections on length of stay. We carried forward the effect of MRSA over time, so once a patient had a MRSA infection, his or her status was yes for all subsequent days until discharge because we expect the effect of an infection to persist; how long the effect persists was one of the questions we addressed in the time-dependent covariate model. The exposures are listed in Table 1.

**Ethics approval**

The hospital ethics committee waived the need for patient consent. It also agreed to use of this anonymized patient database for this study.

**Statistical methods**

Daily ICU data were used for all analyses. Table 2 shows a subset of the data for 2 subjects. Each row corresponds to 1 patient-day in the ICU.

Patient 3154 was admitted to the ICU on February 7 and was discharged on February 10, giving a length of stay of 4 days and a final outcome of “discharged.” Patient 3163 was admitted on February 9 and died on February 12, giving a length of stay of 4 days and a final outcome of “died.” The TISS score (a measure of the level of patient care required) is a time-dependent exposure that changed from day to day (8). In this table, the nominal response variable “outcome” describes each patient’s day-to-day status (stayed, discharged, or died). We assumed that this nominal response had a multinomial distribution, so we examined the probability of “stayed,” “discharged,” and “died” for the i th patient day (i = 1, . . . , n), denoted as πi1, πi2, and πi3, respectively (12, chapter 8). We were interested in the association between these 3 probabilities and the exposures;
The above formulation satisfies the multinomial assumption, hence, we used a nominal logistic regression (or multinomial) model, defined as

\[
\hat{\pi}_{ij} = \frac{1}{1 + \exp(r_{i2} + \exp(r_{i3}))}, \quad i = 1, \ldots, n,
\]

\[
\hat{\pi}_{ij} = \frac{\exp(r_{ij})}{1 + \exp(r_{i2} + \exp(r_{i3}))}, \quad i = 1, \ldots, n, \quad j = 2, 3,
\]

\[
r_{ij} = x_i^T b_j, \quad i = 1, \ldots, n, \quad j = 2, 3,
\]

where \( x_i \) is a set of exposures for row \( i \) of the data and \( b_j \) is the vector of parameters for outcome \( j (b_{j1}, \ldots, b_{jn}) \), where \( j = 2 \) represents ICU discharge and \( j = 3 \) represents death. The above formulation satisfies the multinomial assumption that \( \pi_{i1} + \pi_{i2} + \pi_{i3} = 1 \).

The exponential of \( b_{jk} \) gives the relative risk ratio for a 1-unit increase in the values of covariate \( k \), of being in category \( j = 2, 3 \) relative to category \( j = 1 \) (staying), given that the other covariates are held constant. For example, a value of \( \exp(b_{2,5}) = 2 \) would mean that the relative risk of being discharged would be twice as likely as staying when covariate 5 is increased by 1. A relative risk ratio is similar to an odds ratio but is necessarily more complicated because of the multiple response categories.

This nominal logistic regression model can be thought of as a discrete-time longitudinal survival model (13, section 10.2.3) and is closely related to a competing-risks model (14). To realize this, consider that the multinomial model estimates the probability of death or discharge on each day. Similarly, if we used a competing-risks model, we would estimate the probability of death or discharge in a short period of time. Previous work has shown the similarity between a logistic longitudinal survival model (i.e., \( j = 1, 2 \)) and Cox regression (15). We have extended this similarity by changing the logistic model to a nominal logistic model and the Cox regression to a competing-risks model.

One advantage of using a nominal logistic model with a longitudinal structure is the ability to incorporate random effects, which are useful for modeling heterogeneity and allow the model to account for some of the large, unexplained variation in length of stay. We considered models with a random intercept for each patient admission, allowing the probability of discharge and death to vary between admissions. Thus, regression equation 1 becomes

\[
r_{ij} = x_i^T b_j + z_{ij} I(s_i = s), \quad i = 1, \ldots, n, \quad j = 2, 3,
\]

where \( z_{ij} \) is the random intercept for admission \( s \) for discharge \( (j = 2) \) and death \( (j = 3) \). As before, we do not need to specify an intercept for the reference category \( (j = 1) \). \( I() \) is an indicator function that matches the admission number on row \( i \) of the data to admission \( s \). We used a multivariate
Normal distribution to create each admission’s death and discharge intercept

\[ z_s \sim N(0, \Omega), \quad s = 1, \ldots, m, \]

where \( \Omega \) is a \( 2 \times 2 \) variance-covariance matrix and \( m \) is the total number of admissions.

We also considered models with time-dependent covariates, allowing the effect of MRSA infection on subsequent ICU stay to vary with the number of days in the ICU when infected. To do so, we changed regression equation 1 to

\[ r_{ij} = x'_i \beta_j + x'_i c_{jd} (d_i = d), \quad i = 1, \ldots, n, \quad j = 2, 3, \quad (2) \]

where \( x'_i \) is a time-dependent exposure and \( d_i \) is days since ICU entry for row \( i \) of the data. The parameter \( c_{jd} \) is then the effect of \( x'_i \) on day \( d \) and is estimated separately for discharge (\( j = 2 \)) and death (\( j = 3 \)). We estimated the \( c_{jd} \) parameters as normally distributed random effects given by

\[ c_{jd} \sim N(\mu_c, \sigma^2_c). \]

Because the number of patients not discharged or dead becomes small as \( d \) becomes large, we truncated the time-varying intercept after 21 days. So, for \( d \geq 22, c_{jd} = c_{222} \).

Equation 2 can also be used to model a lagged effect for a time-dependent covariate if \( d \) is defined as number of days since \( x'_i \) equaled \( X \). This lagged effect allows the effect of the covariate to change after a specific event (\( x'_i = X \)). In this analysis, we allowed the effect of MRSA infection to vary with the time since first becoming infected because we were interested in whether the effect of the MRSA infection on discharge waned with increasing time since infection.

**Estimating the excess length of stay and risk of death**

To calculate the excess length of stay due to infection on day \( d \), we subtracted the survivor functions from the day of infection onward using

\[
E(\text{excess length of stay}\mid \text{MRSA on day } d) = \sum_{t=d}^{m} S(t\mid \text{MRSA on day } d) - S(t\mid \text{no MRSA}).
\]

We evaluate the sum up to only some limit \( m \) because, for large values of \( t \), the survivor functions become very small. In this analysis, we use a limit of \( m = 21 \) days. We estimated the survivor function at day \( t \) by multiplying the probabilities of staying from day 1 up to day \( t \),

\[
S(t) = \pi_{j1} (t' = 1, x_j) \times \pi_{j2} (t' = 2, x_j) \times \ldots \times \pi_{jn} (t' = t, x_j),
\]

where \( x_j \) is a set of covariates (we have used a notation different from the one above to emphasize the dependence of the probability on time and the covariates).

To calculate the excess risk of death due to infection, we subtracted the estimated proportion of admissions ending in death for admissions where a MRSA infection occurred from the estimated proportion for admissions where an infection did not occur. As per the estimates of excess length of stay, we specified that the estimate depend on the day of infection. We estimated the excess length of stay and risk of death for an infection occurring during each of the first 21 days in the ICU and for 3 different day 1 APACHE II scores—10, 20, and 30—reflecting a range of morbidity.

**Model fitting and building**

We fitted a number of different models with the model extensions detailed above, using the same set of exposures for each model (Table 1). We selected the best-fitting model using the deviance information criterion. A difference of 10 in the deviance information criterion is considered substantial (http://www.mrc-bsu.cam.ac.uk/bugs/winbugs_dictpage.shtml).

The models were fitted in a Bayesian framework by using the WinBUGS package (16), with vague priors for all unknown parameters. We used a vague Normal prior with a zero mean and a variance of 1,000 for all regression parameters and a gamma prior with a shape and inverse scale parameter of 0.001 for all inverse-variance parameters.

**Missing data**

As shown in Table 1, there were some missing data for the TISS score and just 3 missing scores for the day 1 APACHE II. The TISS scores were most often missing on the day of a patient’s discharge or death. To prevent these important days from being lost from the analysis, we imputed the TISS scores using a random effect for each admission given by

\[
x_{ik} \sim N(\mu_i, \sigma^2_i), \quad i = 1, \ldots, n, \quad k = 1, \ldots, m_i,
\]

where \( m_i \) is the number of days observed for subject \( i \), \( \mu_i \) is the mean score for each subject, and \( \sigma^2_i \) and \( \sigma^2_i \) estimate the between- and within-subject variance, respectively. This imputation was made in WinBUGS in tandem with estimation of the parameters governing discharge and death. To keep the regression model and imputation separate, we used the “cut” function (17).

**RESULTS**

There were 4,569 separate admissions leading to 44,505 days spent in the ICU. Length of patient stays ranged from 0 days to 363 days and were highly skewed, with a mean of 8.8 days (standard deviation, 14.1) and a median of 4 days (interquartile range, 2–11). MRSA was cultured from 864 patients, of whom 335 developed a MRSA infection that led to 6,696 infected-days: 15.0% of the total patient-days in the ICU. One-hundred six (31.6%) of the 335 admitted persons who developed a MRSA infection died compared with 916 of 4,234 (21.6%) of those admitted without a MRSA infection.

We fitted 7 different models to the data, and they are compared using the deviance information criterion in Table 3. Model I had no random intercepts or time-dependent
covariates (refer to Table 1 for a list of exposures). Model II built on model I by including a time-dependent intercept to account for the fact that risk of death or discharge might vary during a patient’s stay in the ICU. Doing so decreased the deviance information criterion by 710, indicating a substantially improved model fit. The probabilities that a patient stayed, died, or was discharged during the first 21 days in the ICU are shown in Figure 1. The most notable feature is the change from a low probability of discharge of 0.08 on the day of admission to the intensive care unit (ICU) (estimates from model II; refer to the text for details). Mean probabilities from model II for a female patient aged 60 years, with a day 1 Acute Physiology and Chronic Health Evaluation II score of 20 and daily therapeutic intervention scoring system score of 40, who is not ventilated or filtrated, in the specialty category of medicine + acute renal failure + cardiology.

Table 3. Comparison of Different Models by Using the Deviance Information Criterion

<table>
<thead>
<tr>
<th>Model</th>
<th>Effective No. of Parameters</th>
<th>Mean Deviance</th>
<th>DIC</th>
<th>Difference in DICa</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: No time-dependent covariates or random intercepts</td>
<td>333.0</td>
<td>27,070</td>
<td>27,400</td>
<td></td>
</tr>
<tr>
<td>II: Model I + time-dependent intercept</td>
<td>382.6</td>
<td>26,310</td>
<td>26,690</td>
<td>710</td>
</tr>
<tr>
<td>III: Model II + random admission intercepts</td>
<td>2,148</td>
<td>24,020</td>
<td>26,170</td>
<td>520</td>
</tr>
<tr>
<td>IV: Model III + time-dependent MRSA</td>
<td>2,117</td>
<td>24,060</td>
<td>26,180</td>
<td>-10</td>
</tr>
<tr>
<td>V: Model III + lagged time-dependent MRSA</td>
<td>2,208</td>
<td>23,930</td>
<td>26,140</td>
<td>30</td>
</tr>
<tr>
<td>VI: Model V – daily TISS + day 1 TISS</td>
<td>1,640</td>
<td>25,370</td>
<td>27,370</td>
<td>-1,230</td>
</tr>
<tr>
<td>VII: Model III – daily TISS</td>
<td>1,649</td>
<td>25,750</td>
<td>27,400</td>
<td>-3,340</td>
</tr>
</tbody>
</table>

Abbreviations: DIC, deviance information criterion; MRSA, methicillin-resistant *Staphylococcus aureus*; TISS, therapeutic intervention scoring system.

a Lower-numbered model minus higher-numbered model. A lower DIC indicates a better model fit; a DIC difference of 10 is considered substantial.

Figure 1. Probability that a patient stayed (dark gray), was discharged (medium gray), or died (light gray), by number of days since admission to the intensive care unit (ICU) (estimates from model II; refer to the text for details). Mean probabilities from model II for a female patient aged 60 years, with a day 1 Acute Physiology and Chronic Health Evaluation II score of 20 and daily therapeutic intervention scoring system score of 40, who is not ventilated or filtrated, in the specialty category of medicine + acute renal failure + cardiology.

Models I–V all adjusted for the daily TISS score. We were concerned that TISS score could be affected not only by the underlying severity of patient illness but also by MRSA infection. If so, adjusting for the daily TISS score would bias the estimate of the effect of MRSA on death and discharge (18), possibly causing the model to miss a true association between MRSA and death. We therefore fitted 2 more models (model VI and VII). Model VI was the same as model V but adjusted for day 1 TISS only instead of daily TISS. The effect of MRSA infection on death changed little, but the relative risk ratio of discharge strengthened to 0.69 (95% CI: 0.55, 0.84). Model VII was the same as model III
but without any adjustment for TISS score; the effect on discharge was similar (RRR = 0.68, 95% CI: 0.54, 0.82).

The risk ratios of death and discharge relative to staying, from model VII, are shown in Table 4. The strongest reduction in the risk of discharge was associated with ventilation, whereas hemofiltration was associated with the strongest increase in the risk of death. There was little evidence of difference in the risks of discharge or death between the 5 specialty categories. The risk of discharge varied by day of the week, being significantly lower on the weekend compared with Wednesday.

Figure 4 shows the relative risk ratio of discharge and death after a MRSA infection for the 7 different models. For models IV–VI, the effect plotted is the mean over all times. Including a time-dependent intercept greatly changed the effect of MRSA infection, as shown by the differences in relative risk ratios between models I and II. Adding a random intercept meant that the effect of MRSA significantly decreased the risk of discharge (model III compared with model II). The mean relative risk ratios for models IV and V were similar to those for model III but with slightly wider credible intervals. Models VI and VII showed the strongest reduction in the risk of discharge after MRSA infection.

Figure 5 shows the mean excess length of stay and risk of death due to a MRSA infection according to day 1 APACHE scores. For an infection on day 0, a patient with an APACHE II score of 10 would have a mean extra length of stay of 0.3 days (95% CI: 0.1, 0.5). A sicker patient with an APACHE II score of 30 would have a longer stay of 1.2 days (95% CI: 0.5, 2.0). The increased risk of death is as high as 0.13 (95% CI: 0.06, 0.21) for a patient with an APACHE II score of 30 who becomes infected on day 1. The extra risk of death declines linearly with increasing time to infection and is much lower for patients with a lower APACHE II score.

It is possible that our definition of a MRSA infection included patients who had colonization only. We therefore

Figure 2. Daily mean relative risk ratio (RRR; solid lines) and 95% credible interval (dashed lines) of discharge from the intensive care unit (left panel) and of death (right panel) relative to staying after developing a methicillin-resistant Staphylococcus aureus (MRSA) infection, dependent on the time delay between admission and infection (estimates from model IV; refer to the text for details).

Figure 3. Daily mean relative risk ratio (RRR; solid lines) and 95% credible interval (dashed lines) of discharge from the intensive care unit (left panel) and of death (right panel) relative to staying after developing a methicillin-resistant Staphylococcus aureus (MRSA) infection, by number of days since infection (estimates from model V; refer to the text for details).
credible intervals (RRR of discharge and death was slightly stronger but with wider criteria using model VII. The effect of infection on both with an infection from 335 to 105. We tested the new MRSA criteria). Doing so reduced the overall number of patients otic start and systemic inflammatory response syndrome tures or other sterile-site aspirates (together with an antibi-
ated using only positive blood cultures or other sterile-site aspirates (together with an antibiotic start and systemic inflammatory response syndrome criteria). Doing so reduced the overall number of patients with an infection from 335 to 105. We tested the new MRSA criteria using model VII. The effect of infection on both discharge and death was slightly stronger but with wider credible intervals (RRR of discharge = 0.60, 95% CI: 0.41, 0.83; RRR of death = 1.25, 95% CI: 0.82, 1.80). We might expect to see a stronger association if the new criteria more accurately captured true infections. Widening of the credible intervals is likely due to a reduction in power because of the smaller number of infections.

Table 4. Relative Risk Ratios and 95% Credible Intervals of Daily Discharge and of Death Relative to Staying in the Intensive Care Unit

<table>
<thead>
<tr>
<th>Variable (Unit)</th>
<th>Discharge</th>
<th>Death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RRR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Age (10-year increase)</td>
<td>0.93</td>
<td>0.90, 0.96</td>
</tr>
<tr>
<td>Sex (male vs. female)</td>
<td>1.12</td>
<td>1.01, 1.25</td>
</tr>
<tr>
<td>Day 1 APACHE II score</td>
<td>0.66</td>
<td>0.63, 0.69</td>
</tr>
<tr>
<td>Ventilation (yes vs. no)</td>
<td>0.05</td>
<td>0.04, 0.06</td>
</tr>
<tr>
<td>Hemofiltration (yes vs. no)</td>
<td>0.30</td>
<td>0.25, 0.36</td>
</tr>
<tr>
<td>MRSA infection (yes vs. no)</td>
<td>0.68</td>
<td>0.54, 0.82</td>
</tr>
<tr>
<td>Specialty (surgery)</td>
<td>0.87</td>
<td>0.74, 1.01</td>
</tr>
<tr>
<td>Specialty (cardiothoracic surgery)</td>
<td>0.94</td>
<td>0.80, 1.09</td>
</tr>
<tr>
<td>Specialty (orthopedics)</td>
<td>0.84</td>
<td>0.55, 1.24</td>
</tr>
<tr>
<td>Specialty (intensive therapy unit referrals)</td>
<td>0.93</td>
<td>0.78, 1.11</td>
</tr>
<tr>
<td>Day of the week (Monday)</td>
<td>0.89</td>
<td>0.77, 1.05</td>
</tr>
<tr>
<td>Day of the week (Tuesday)</td>
<td>0.94</td>
<td>0.81, 1.11</td>
</tr>
<tr>
<td>Day of the week (Thursday)</td>
<td>1.00</td>
<td>0.85, 1.17</td>
</tr>
<tr>
<td>Day of the week (Friday)</td>
<td>1.06</td>
<td>0.92, 1.24</td>
</tr>
<tr>
<td>Day of the week (Saturday)</td>
<td>0.72</td>
<td>0.62, 0.86</td>
</tr>
<tr>
<td>Day of the week (Sunday)</td>
<td>0.63</td>
<td>0.53, 0.75</td>
</tr>
</tbody>
</table>

Abbreviations: APACHE, Acute Physiology and Chronic Health Evaluation; CI, credible interval; MRSA, methicillin-resistant Staphylococcus aureus; RRR, relative risk ratio.

We found that, in all 7 multinomial models considered, MRSA infection decreased the risk of discharge. Although the magnitude of this effect varied considerably between models, the 3 models that gave (by some margin) the best fit to the data (III, IV, and V) all yielded remarkably similar estimates: patients with MRSA infections had a relative risk of discharge (compared with staying) that was about 20% lower than that for patients without MRSA infections.

Additional stay attributed to the MRSA infection was found to be longer for sicker patients (as measured by the APACHE II score) and for infections occurring earlier during the ICU stay (Figure 5). The mean excesses are for all admissions, so infections occurring later have a much smaller attributable cohort and so cause fewer overall excess stays. Sicker patients have less physical reserve and may be less able to cope with an infection, hence their increased length of stay after infection compared with healthier patients.

One potential drawback of our approach is the retrospective nature of the infection diagnosis based on systemic inflammatory response syndrome criteria, antibiotic start, and culture of MRSA from a clinical site. This approach led to a diagnosis of MRSA infection in 335 (39%) of 864 patients colonized with MRSA, which, although high, is likely explained by the particular virulence of MRSA compared with other hospital bacteria and the hyperinvasive nature of the variant of MRSA that was circulating in the ICU at that time (9). Use of a sterile-site definition, which does not have the potential to inadvertently include colonized patients, resulted in only a slightly stronger association between infection and reduction in discharge and increase in death, supporting the view that the wider definition of infection we used was reasonably accurate.

We assumed that once patients were infected, they remained infected for the rest of their stay rather than returning.
to an uninfected state, an assumption that gains some support from evidence that the effect of infection remained constant in the model (Figure 3). An alternative assumption would be that patients can return to the uninfected state and therefore can have repeat infections. When a 2-week cutoff was used before allowing patients to return to the uninfected state, which equates to a standard 2-week treatment course for serious infections in the ICU, only 11% of all infected patients or 4% of sterile-site-infected patients went on to have further infections. This finding suggests that this alternative assumption would not have significantly altered our analysis.

Our approach to estimating additional length of stay caused by hospital infections overcomes the pitfalls that affect much of the literature and should therefore provide more reliable estimates. However, an alternative analytical approach would have been to use a multistate model (14, 19, 20).

Such multistate models can be used to model the flow of patients through a set of defined states. For example, patients may start in the “hospital entry” state, and some may subsequently move to an “infected” state and then to a “discharged” state. However, an important advantage of the approach we used is the relative ease of incorporating lagged covariates and random effects. Nonetheless, the 2 approaches have similarities, and the multinomial model we used can be thought of as a discrete-time analogue of a multistate model (when comparable models could be fitted, they were in fact found to give very similar results). One potential drawback of the multinomial model is that the data must be equally spaced, and, in our case, all times were rounded to the nearest day. However, the model could be readily extended to smaller time steps if required and if data were available.

Another advantage of multistate models is their ability to condition on final outcome, so that the excess length of stay is different for patients who die compared with patients who are discharged (21). The extra morbidity of MRSA infection may lead to a shorter stay (faster death) for very frail patients but a longer stay for other patients.

The aim of this study was to estimate the degree to which MRSA infection causes increased length of stay, rather than to simply document associations. This goal raises important issues about identifying confounders. Controlling for the time-dependent exposure of daily TISS score gave a better fit to the data (Table 3), but it is also likely that MRSA infection will affect the TISS score, particularly by prompting new antibiotic starts, catheter insertions, and treatment with vasoactive agents. Therefore, adjusting for the daily TISS score in the model would not only be unnecessary but also be potentially actively harmful and introduce bias (18). This consideration motivated model VI (which includes only the day 1 rather than the daily TISS score) and model VII (without any TISS score), which are therefore not vulnerable to this problem. These models showed an increased effect of MRSA infection on discharge compared with the other models with time-dependent covariates (Figure 4) (18).

Being able to compare the fit of the models by using the deviance information criterion is one of the advantages of a Bayesian framework (Table 3). One surprising result was the reduction in the effective number of parameters from 2,208 for model V to 1,604 for model VI, when the only change was fitting TISS as a time-independent, instead of time-dependent, covariate (hence, we might have expected a reduction of about 40 parameters). The large reduction in the effective number of parameters is due to the day 1 TISS score explaining much of the between-admission variation in the risk of discharge and death. This variation is no longer modeled by the random subject intercept, so many fewer parameters are needed. Nonetheless, despite the poorer fit to data as measured by the deviance information criterion, knowledge of the likely causal pathways suggests that models VI and VII would give the most reliable estimate of the impact of MRSA infection on length of stay in an ICU.
ACKNOWLEDGMENTS

Author affiliations: School of Public Health and Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, Queensland, Australia (Adrian G. Barnett, Nicholas Graves); Statistics, Modelling, and Bioinformatics Department, Centre for Infections, Health Protection Agency, London, United Kingdom (Ben Cooper, Julie Robotham); Department of Infection, Guy’s and St. Thomas’ National Health Service Foundation Trust, London, United Kingdom (Rahul Batra, Jonathan Edgeworth); and Department of Infectious Diseases, King’s College London School of Medicine at Guy’s King’s College and St. Thomas’ Hospitals, London, United Kingdom (Jonathan Edgeworth).

J. E. and R. B. received funding under the Department of Health’s Healthcare-Associated Infections Initiative and Guy’s and St. Thomas’ Charity. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health. The authors acknowledge financial support from the Department of Health via the National Institute for Health Research comprehensive Biomedical Research Centre award to Guy’s and St. Thomas’ National Health Service Foundation Trust in partnership with King’s College London and King’s College Hospital National Health Service Foundation Trust.

Computational resources and services used in this work were provided by the High Performance Computer and Research Support Unit, Queensland University of Technology, Brisbane, Australia.

The authors thank Dr. Martin Wolkewitz for his useful comments on a draft of this paper.

Conflict of interest: none declared.

REFERENCES