Appendix

Polynomial Distributed Lag (PDL) model

A PDL model was built to detect and quantify the lagged effects of antimicrobial use on the % methicillin-resistant *Staphylococcus aureus* (MRSA). In a PDL model, the relationship between the dependent variable (resistance) and the independent variables (past resistance and antimicrobial use) should evolve smoothly over time, through the use of "polynomial lags." The optimum PDL model was arrived at by the "general-to-specific" econometric methodologic characteristics. This meant that, initially, many possible independent variables were included in the model, some of which were ultimately found to be irrelevant. Additionally, for all the independent variables, lags of up to 8 months were initially included to identify direct effects. The initial dynamic regression model with PDLs considering %MRSA series as the dependent variable and several antimicrobial drug use series as explanatory series was the following:

$$\%MRSA_t = \alpha + \sum_{i=1}^{8} \beta_{0i} \%MRSA_{t-i} + \sum_{i=0}^{8} \beta_{1i} MAC_{t-i} + \sum_{i=0}^{8} \beta_{2i} 3GC_{t-i} + \sum_{i=0}^{8} \beta_{3i} FQ_{t-i} + \sum_{i=0}^{8} \beta_{4i} PIB_{t-i} + \varepsilon_t$$

with PDL restrictions on the coefficients of antimicrobial use and where MAC means macrolide use, 3GC third-generation cephalosporin use, FQ fluoroquinolone use and PIB use of penicillins with ß-lactamase inhibitors. The model was initially estimated on the full study period, i.e., January 1996–December 2000, using a degree q_j of the polynomial equal to 3. The estimated model was compatible with normal white noise errors (absence of autocorrelation and absence of heteroskedasticity), and no signs of nonmodeled nonlinearities were seen.

This initial model was then simplified to eliminate irrelevant antimicrobial drug uses and unnecessary lags. In the first steps of the simplification, all antimicrobial drugs were kept in the model, and the simplification took the form of reducing the order of the polynomial and eliminating unnecessary lags. Along this process, use of penicillins with ß-lactamase inhibitors did not appear to play a significant role and was eliminated from the model. We also tried to introduce use of each of the other antimicrobial drug classes that showed a relationship in Table 3; however, none appeared to play an important role, and they were not included in the model. Further simplification of the distributed lags of macrolide use, third-generation cephalosporin use, and fluoroquinolone use of the %MRSA itself led to a model in which, through CUSUM and CUSUMSQ statistics, a structural change was detected around the middle of 1997. Application of the Chow test located the change in June 1997. The %MRSA was virtually zero in 1996 and started to increase at the beginning of 1997, which was consistent with the fact that the MRSA epidemic strain, resistant to macrolides and fluoroquinolones, only became predominant in 1997. In 1996, 56% and 50% of MRSA isolates were resistant to erythromycin and ciprofloxacin, respectively, whereas these percentages suddenly rose to 92% and 89%, respectively, in 1997. Data before June 1997 were considered as not being part of the outbreak and were therefore not included in the final model. The validity of the simplified, final model from June 1997 onwards was checked by a battery of specification and diagnostic tests to verify the absence of autocorrelation of residuals, absence of heteroskedasticity, normality of residuals, absence of nonmodeled nonlinearities and absence of structural change.

The basic measure of forecasting quality, Root Mean Squared Error of Forecast (RMSEF) was also computed, which provided an average measurement of the amount by which the model over- or underestimated the %MRSA. RMSEF was calculated for a model without antimicrobial drug use (based on past %MRSA only) and compared with that of the final model, which included antimicrobial drug use.