

Supplementary material 2

Factors associated with cUTI-related healthcare costs – statistical details

The analysis of the factors associated with cUTI-related healthcare costs was undertaken using multivariate regression analysis using patient level cost data. The dependent variable was total cost per patient estimated as described above.

To account for skewness of the cost data, generalised linear models with gamma family and log link were used [1]. We also considered using log Normal, Gaussian, inverse Gaussian and negative binomial distributions, but the gamma model gave the best fit in terms of the Akaike Information Criterion. We did not include as explanatory variables any of the variables used to construct the total cost per patient. We also exclude variables with a high collinearity ($r > 0.6$). The variable selection in the reduced model was undertaken using forward and backward inclusion methods. P-values below the 5% level are regarded as statistically significant. Values between 5 and 10% are regarded as weakly significant.

For the quantitative interpretation of the effect of each variable, we computed marginal effects at the mean values of the included covariates. The impact of unobserved heterogeneity due to the hierarchical structure of the data is explored and accounted for by considering country fixed effects models. We also adjust for clustering at the site level by computing robust standard errors, and control for the patient episode number by including this indicator as an explanatory variable in the models.

References

1. Barber J, Thompson S. Multiple regression of cost data: use of generalised linear models. *J Health Serv Res Policy* 2004;**9**:197-204.
2. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG et al. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria:

an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infec.* 2012; **18**:268-81