Revealing antibiotic cross-resistance patterns in hospitalized patients through Bayesian network modelling

Stacey S Cherny[1,2], Daniel Nevo[3], Avi Baraz[1,2,3], Shoham Baruch[1,2], Ohad Lewin-Epstein[4], Gideon Y Stein[5,6], Uri Obolski[1,2,*]

[1]School of Public Health, Tel Aviv University, Tel Aviv, Israel

[2]Porter School of the Environment and Earth Sciences, Tel Aviv University, Tel Aviv, Israel

[3]Department of Statistics and Operations Research, Tel Aviv University, Tel Aviv, Israel

[4]Department of Molecular Biology and Ecology of Plants, Tel Aviv University, Tel Aviv, Israel

[5]Internal Medicine “A”, Meir Medical Center, Kfar Saba, Israel

[6]Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel

[*]to whom correspondence should be addressed: uriobols@tauex.tau.ac.il

Keywords: antibiotic resistance, Bayesian network modelling, cross-resistance, statistical modelling

Abstract

Objectives:
Microbial resistance exhibits dependency patterns between different antibiotics, termed cross-resistance and cross-sensitivity. These patterns differ between experimental and clinical settings. It is unclear whether the differences result from biological reasons or from confounding, biasing results found in clinical settings. We set out to elucidate the underlying dependency patterns between resistance to different antibiotics from clinical data, while accounting for patient characteristics and previous antibiotic usage.
Methods: Additive Bayesian network modelling was employed to simultaneously estimate relationships between variables in a dataset of bacterial cultures derived from hospitalized patients and tested for resistance to multiple antibiotics. Data contained resistance results, patient demographics, and previous antibiotic usage, for five bacterial species: *E. coli* (n=1054), *K. pneumoniae* (n=664), *P. aeruginosa* (n=571), CoNS (n=495), and *P. mirabilis* (n=415).
Results: All links between resistance to the various antibiotics were positive. Multiple direct links between resistance of antibiotics from different classes were observed across bacterial species. For example, resistance to gentamicin in *E.coli* was directly linked with resistance to ciprofloxacin (OR = 8.39, 95%CI[5.58, 13.30]) and sulfamethoxazole-trimethoprim (OR = 2.95, 95%CI[1,97, 4.51]). In addition, resistance to various antibiotics was directly linked with previous antibiotic usage.
Conclusions: Robust relationships among resistance to antibiotics belonging to different classes, as well as resistance being linked to having taken antibiotics of a different class, exist even when taking into account multiple covariate dependencies. These relationships could help inform choices of antibiotic treatment in clinical settings.

Introduction

The increase in antibiotic resistance frequencies of bacterial pathogens is leading to an ever-growing burden on healthcare systems, in terms of costs for newer drugs, extended hospital stays, and increased followup visits, as well as increased cost in terms of lost productivity and mortality [1,2]. According to a new report by the United States Centers for Disease Control, annually in the USA, 2.8 million people suffer from infections that are resistant to antibiotics, resulting in 35,000 deaths per year. They reported that multidrug-resistant *P. aeruginosa*, one of the bacterial species examined in the present paper, was estimated to cause nearly 10% of the deaths attributable to antibiotic resistant bacteria and over $750M dollars in direct healthcare costs in the USA in 2017 [3]. Hence, understanding the nature of the relationships among patient covariates and the resistance to the different drugs in the current arsenal of antibiotics is crucial.

Beyond well studied, patient-related predictors for antibiotic resistance, such as age, patients' independence status, previous antibiotic usage, and more [4–6], resistance to different antibiotics is often not independent. Cross-resistance, for instance, refers to the existence of positive associations between resistance to different antibiotics across bacterial pathogens. Such dependencies are a phenomenon known for nearly as long as antibiotics have been available [7], with the mechanisms behind it numerous [8].

Cross-resistance can occur due to inherent biological attributes of bacterial pathogens, when antibiotics have similar mechanisms of action and hence shared mechanisms of resistance; cross-resistance between chemically dissimilar antibiotics can also occur due to horizontal gene transfer of genetic elements coding for resistance to multiple antibiotics [9]. Conversely, there is also evidence of cross-sensitivity, or collateral-sensitivity, where negative associations between resistance to different antibiotics are observed. However, while cross-sensitivity has been demonstrated in laboratory experiments [10–13], when examining cultures obtained from patients in a hospital setting, it appears that host factors induce mostly positive correlations between resistance to different antibiotics [8]. As a result of the complex relationships of resistance among different antibiotics, examining resistance to a single antibiotic or even pairs of antibiotics is insufficient. A multivariate approach which allows multiple drug resistances as dependent variables is warranted to allow uncovering the underlying structure of the observed resistance patterns, which can include conditional dependencies between the antibiotics. One such approach, that we utilize in this paper, is additive Bayesian network (ABN) modelling [14].

ABN modelling is a purely data-driven approach to inferring underlying probabilistic structure of a set of variables. By essentially searching all possible directed acyclic graphs (DAGs) linking a set of variables, evidence for potential causal links can be revealed from the data without making strong prior assumptions. An approach with few prior assumptions regarding dependencies is of great utility in this area, where resistance to many different drugs and the presence of several important covariates need to be analyzed, but no strong hypotheses regarding many of the connections are available.The method has been successfully applied in veterinary studies of disease [15], as well as to antibiotic resistance in animals [16–18]. A

related Bayesian approach was employed in a study of meticillin-resistant *Staphylococcus aureus* transmission in humans [19]. However, to our knowledge the present application is the first time ABN has been used in a study of antibiotic resistance to infer cross-resistance patterns in human infections.

In this study, we employed ABN modelling to explore the interrelationships among bacterial resistance to several antibiotics, using data from bacterial cultures obtained from a single hospital in Israel. We used ABN modeling separately on each of the five bacterial species most frequent in our dataset, to examine predictors of resistance and the interrelationships among the antibiotics tested.

Materials and methods

Ethics

The study was approved by the Helsinki Committee of Rabin Medical Center.

Data

We obtained data pertaining to bacterial cultures drawn in Rabin Medical Center, Israel, from 2013-05-01 to 2015-12-31. The corresponding demographics, previous hospitalizations, and previous in-hospital antibiotic usage in the year prior to the infection, of patients from whom the cultures were drawn, were also available. Bacterial cultures were tested for antibiotic resistance for an array of antibiotics which had varying rates of resistance, and results of non-susceptibility and resistance were combined into a 'resistant' category. Bacterial infections were considered nosocomial if infections occurred >48 hours after admission. A summary of these variables is presented in Table 1. For our analysis, we selected the five bacterial species with the largest sample sizes available in the dataset (in order of decreasing sample size): *Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa, Proteus mirabilis,* and Coagulase-negative *staphylococci* (CoNS).

Table 1: Descriptive statistics of patients and bacterial isolates.

	E. coli **(n=1054)**	***K. pneumoniae*** **(n=664)**	***P. aeruginosa*** **(n=571)**	***P. mirabilis*** **(n=415)**	**CoNS (n=495)**
Age: Mean (SD)	76.7 (14.8)	73.5 (15.0)	71.7 (16.1)	74.8 (14.1)	72.8 (18.0)
Female	568 (53.9%)	251 (37.8%)	197 (34.5%)	131 (31.6%)	221 (44.6%)
T.F.aminoglycoside	142 (13.5%)	137 (20.6%)	120 (21.0%)	101 (24.3%)	90 (18.2%)
T.F.fluoroquinolone	268 (25.4%)	214 (32.2%)	142 (24.9%)	97 (23.4%)	136 (27.5%)
T.F.betalactam	515 (48.9%)	468 (70.5%)	437 (76.5%)	271 (65.3%)	305 (61.6%)
T.F.other	406 (38.5%)	354 (53.3%)	380 (66.5%)	237 (57.1%)	268 (54.1%)
Hosp: Mean (SD)	2.05 (1.29)	2.68 (1.10)	2.91 (1.11)	2.91 (1.29)	2.41 (1.19)

Nosocomial	237 (22.5%)	244 (36.7%)	344 (60.2%)	198 (47.7%)	196 (39.6%)
Polymicrobial culture	192 (18.2%)	225 (33.9%)	289 (50.6%)	213 (51.3%)	35 (7.1%)
Amikacin		93 (14.0%)			
Ampicillin	858 (81.4%)				
Ceftazidime			74 (13.0%)		
Cefuroxime				233 (56.1%)	
Chloramphenicol					239 (48.3%)
Ciprofloxacine	523 (49.6%)	389 (58.6%)		241 (58.1%)	
Erythromycin					328 (66.3%)
Fusidic acid					218 (44.0%)
Gentamicin	216 (20.5%)	279 (42.0%)	87 (15.2%)	219 (52.8%)	196 (39.6%)
Imipenem			91 (15.9%)		
Ofloxacin					291 (58.8%)
Oxacillin					366 (73.9%)
Sulf-Trim	534 (50.7%)			279 (67.2%)	253 (51.1%)

Note: T.F. prefix denotes having taken an antibiotic from the given class; Sulf-Trim = sulfamethoxazole-trimethoprim. Antibiotics listed refer to resistance to those antibiotics. Hosp = Log of the number of days hospitalized in the prior year plus 1.

Statistical analysis

We selected which antibiotics to include in the analysis by keeping only those with minimal missing data and which did not reduce the number of complete cases appreciably (<10% loss). We performed some variable selection to assure stable statistical models with no perfect or near-perfect separation, by not including perfectly or near perfectly correlated antibiotics and selecting only antibiotics which contained a minimum of 3% resistance in each bacterial subsample. Variables excluded from analysis are presented in Supplementary Table S1, along with their tetrachoric correlations with the relevant included variables. This resulted in analysis of between three and seven antibiotics for the five bacterial species, each analyzed separately. When constructing the ABN, the following covariates were included, in addition to the antibiotic resistance variables: demographic variables (age, sex, and days hospitalized in the previous year), binary culture type variables (nosocomial and polymicrobial), and antibiotic use in the previous year. Antibiotics used were first grouped into classes and the three most frequent classes across the entire dataset were included, along with whether any other antibiotics were taken which did not belong to the three largest classes. The classes were aminoglycosides, fluoroquinolones, betalactams, and other antibiotics not part of these three classes. Table 1 presents summary statistics for all variables used in our models.

Data were analyzed using additive Bayesian network (ABN) modelling [14,20], with version 2.1 of the R package abn [21] on an R 3.6.1 installation [22]. Briefly, ABN modelling is a purely data-driven, exploratory approach, originating in machine learning, that is often used

for hypothesis generation for causation among a set of variables. By essentially searching all possible directed acyclic graphs (DAGs) linking a set of variables, a model for the dependency of the variables in the data can be inferred without making strong prior assumptions. This model, depicted in a DAG, shows which variables are directly connected, or linked, via arcs, and the coefficients of these arcs are directly analogous to the adjusted odds ratios obtained from multiple logistic regression analysis. While no assumptions of causal relationships are required, we restricted the model space to disallow causal paths that made no sense for our study: (1) sex could only cause other variables and not be caused by any other variable, (2) age at testing could be caused by sex but no other variable, (3) the four variables for antibiotics given in the previous year could only be caused by age and sex and no other variable, (4) hospitalization in the prior year could not be caused by presence of nosocomial infection but could be caused by any other variable, (5) nosocomial infection could not be caused by resistance to any drug but can be caused by any other variable; and (6) polymicrobial cultures as well as all drug resistances could be caused by any of the variables. In addition, no causal paths were forced to be in the model.

There are several steps involved when implementing an ABN analysis. We first used the exact search method to determine the maximum number of parents (i.e., nodes that have an arc causing another node) needed for any child (i.e., variable caused by another variable) in the model, above which there was no improvement in likelihood, by repeatedly running the exact search and setting the maximum number of parents between one and six. Never were more than four parents needed (see Supplementary Table S2). Next, we used a parametric bootstrapping approach to correct for overfitting [23]. This was done by simulating 1000 samples of the same size as in the original data, by feeding the ABN model chosen (which arcs were present and their parameter estimates) to the JAGS software, version 4.3.0 [24]. This produces random samples generated from the parameter estimates of the chosen model. These samples were each analyzed with ABN in exactly the same way as performed initially. Next, we examined each of the 1000 ABN models and retained only those arcs present in a majority (at least 50%) of them, a common cutoff as suggested by the ABN developers [14]. Supplementary Table S2 contains the number of arcs present in the model both before and after bootstrapping. Finally, we ran ABN on the original dataset constraining the model to only include the selected consensus arcs.This final model was used to produce 95% Bayesian credible intervals for each of the parameter estimates, which are analogous to coefficients from logistic regression models, i.e., the odds ratio (OR) of the effect of parent (independent) on child (dependent) variable. Due to the existence of topologies that are equivalent in terms of likelihood in Bayesian networks [18,25–30], we present the models' arcs and their corresponding parameters, but not interpreting their direction.

Results

Before fitting the ABNs, we removed antibiotic variables which were nearly identical in their resistance patterns. This yielded expected results between similar antibiotics but also some extremely high associations between antibiotics of different classes. For example, we found that in all cultures where they were tested, the betalactams ceftazidime, ceftriaxone, and cefuroxime were all perfectly inter-correlated, and ertapenem and piperacillin were nearly

perfectly correlated with them. However, amikacin and ampicillin were nearly perfectly correlated in the *E.coli* dataset, yet are members of different drug classes. These and additional highly correlated pairs of antibiotics can be found in Table S1. It is important to note that the results derived below are practically identical for members of the antibiotic pairs presented in Table S1.

We present the final ABNs estimated from our data, for five bacterial species, in Figures 1 through 5. Each arc in the plots represents an estimated direct link between the variables, where adjacent numbers are the parameter estimates of the OR of this link. In addition, these values along with their 95% credible intervals are presented in Tables 2-6. Notably, none of the credible intervals contained the value one, i.e. no association, although this is not a required condition by the Bayesian network structure discovery algorithm [14]. This strengthens the notion that the connections found by our models are robust.

To summarize the analyses across the five bacterial species examined, we constructed a matrix of all possible arcs present in any of the five models (see Supplementary Figure 1). Within each cell, we present the number of models which contained the corresponding arc over the number of analyses which could have contained that arc (that is, counting datasets in which both variables were present). For ease of interpretation, variables which appeared in only a single model were omitted (Figure 6). Importantly, the signs of all coefficients, when arcs were found, were consistent throughout the analysis of all five datasets. That is, no link between variables was found negative for a certain bacterial species and positive in another.

The most important results from our analyses were the direct connections found between resistance to different antibiotics. **These were consistently positive throughout all antibiotics and all bacterial species (Tables 2-6)**. This implies that even when accounting for the possible connections between the different variables, acquisition of resistance to a certain antibiotic only served to enhance the probability of resistance to other antibiotics in our data.

Interestingly, we found several robust links between antibiotics of different classes. Direct links from different antibiotic classes to gentamicin were especially prevalent. Gentamicin and sulfamethoxazole-trimethoprim were directly linked in the *E. coli* (Table 2 and Figure 1) and *P. mirabilis* datasets (Table 5 and Figure 4). Furthermore, gentamicin resistance was directly linked to resistance to several betalactam antibiotics. It was directly linked to both resistance to ceftazidime and imipenem (Table 4 and Figure 3) in the *P. aeruginosa* dataset and to cefuroxime (Table 5 and Figure 4) in the *P. Mirabilis* dataset. Gentamicin resistance was also directly linked to resistance to ciprofloxacin, a fluoroquinolone, in the *E. coli, K. pneumoniae,* and *P. mirabilis* datasets. Additionally, sulfamethoxazole-trimethoprim, was directly linked to ciprofloxacin in both datasets where they were tested (*E. coli* [Table 2 and Figure 1] and *P. mirabilis* [Table 5 and Figure 4]).

Figure 1: ABN model for *E. coli*. Arcs denote direct links between variables and adjacent values are ORs derived from the model (see Table 2). Nodes with antibiotic names denote resistance to those antibiotics (enclosed in a green rectangle). T.F. prefix denote variables indicating having taken an antibiotic from the given class in the prior year (enclosed in a red rectangle); Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; AMP = ampicillin; CIP = ciprofloxacin; GEN = gentamicin; Sulf-Trim = sulfamethoxazole-trimethoprim. Sex=male.

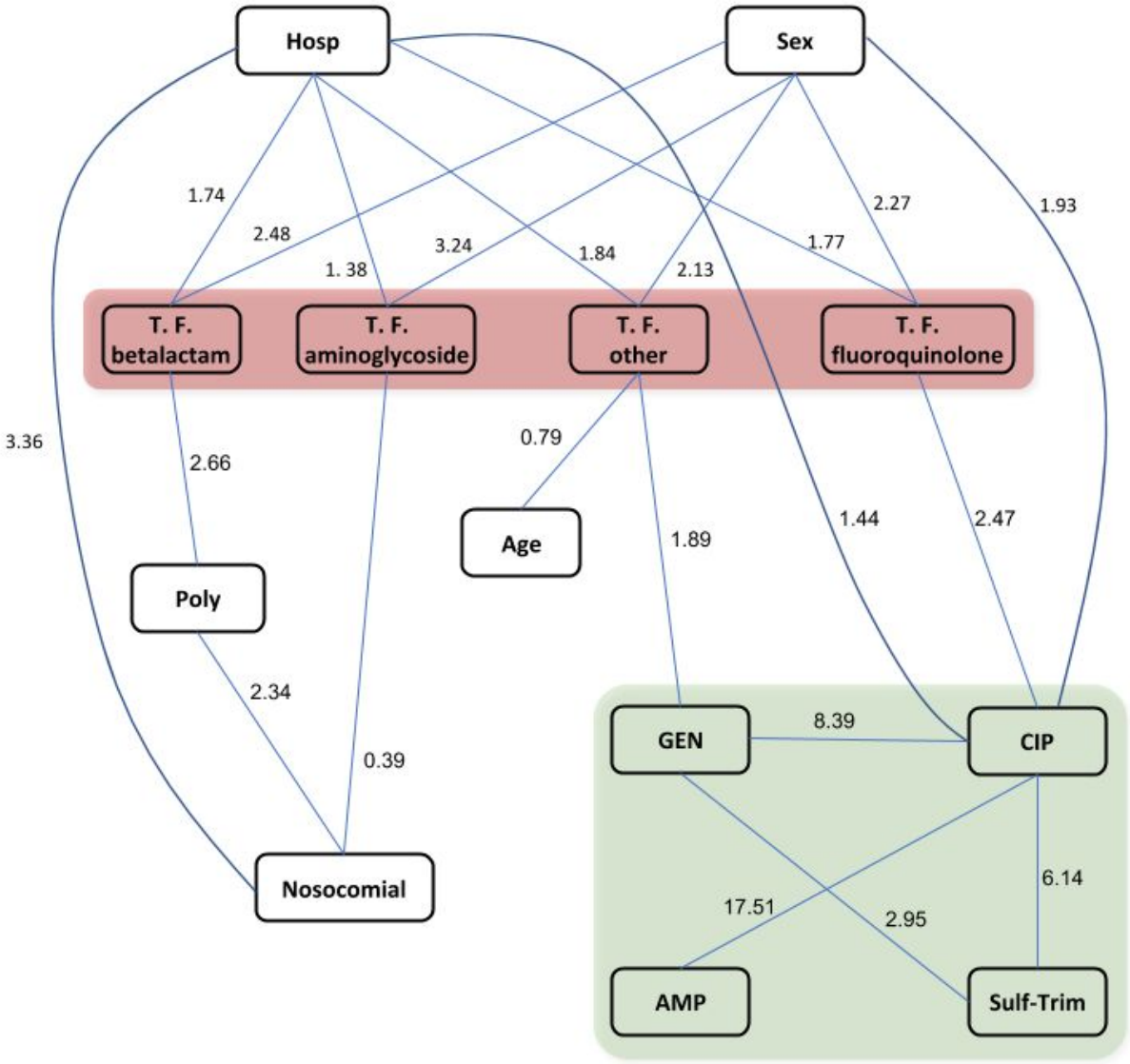


Table 2: *Escherichia coli* OR estimates and their 95% credible intervals. None of the credible intervals contains zero.

Variable|Affected by	2.5%	OR	97.5%
T.F.aminoglycoside|Sex	2.18	3.24	4.79
T.F.fluoroquinolone|Sex	1.69	2.27	3.04
T.F.betalactam|Sex	1.92	2.48	3.18
T.F.other|Age	0.69	0.79	0.90
T.F.other|Sex	1.64	2.13	2.76
Hosp|T.F.aminoglycoside	1.20	1.38	1.58
Hosp|T.F.fluoroquinolone	1.57	1.77	1.98
Hosp|T.F.betalactam	1.56	1.74	1.94
Hosp|T.F.other	1.64	1.84	2.06
Nosocomial|T.F.aminoglycoside	0.23	0.39	0.60
Nosocomial|Hosp	2.75	3.36	4.17
Poly|T.F.betalactam	1.87	2.66	3.82
Poly|Nosocomial	1.64	2.34	3.30
AMP|CIP	10.45	17.51	32.46
CIP|Sex	1.46	1.93	2.51
CIP|T.F.fluoroquinolone	1.72	2.47	3.55
CIP|Hosp	1.23	1.44	1.67
GEN|T.F.other	1.34	1.89	2.64
GEN|CIP	5.58	8.39	13.30
Sulf-Trim|CIP	4.58	6.14	8.19
Sulf-Trim|GEN	1.97	2.95	4.51

Note: T.F. prefix denotes having taken an antibiotic from the given class; Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; AMP = ampicillin; CIP = ciprofloxacin; GEN = gentamicin; Sulf-Trim = sulfamethoxazole-trimethoprim.

Figure 2: ABN model for *K. pneumoniae.* Arcs denote direct links between variables, and adjacent values are ORs derived from the model (see Table 3). Nodes with antibiotic names denote resistance to those antibiotics (enclosed in a green rectangle). T.F. prefix denotes variables indicating having taken an antibiotic from the given class in the prior year (enclosed in a red rectangle); Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; AMK = amikacin; CIP = ciprofloxacin; GEN = gentamicin.

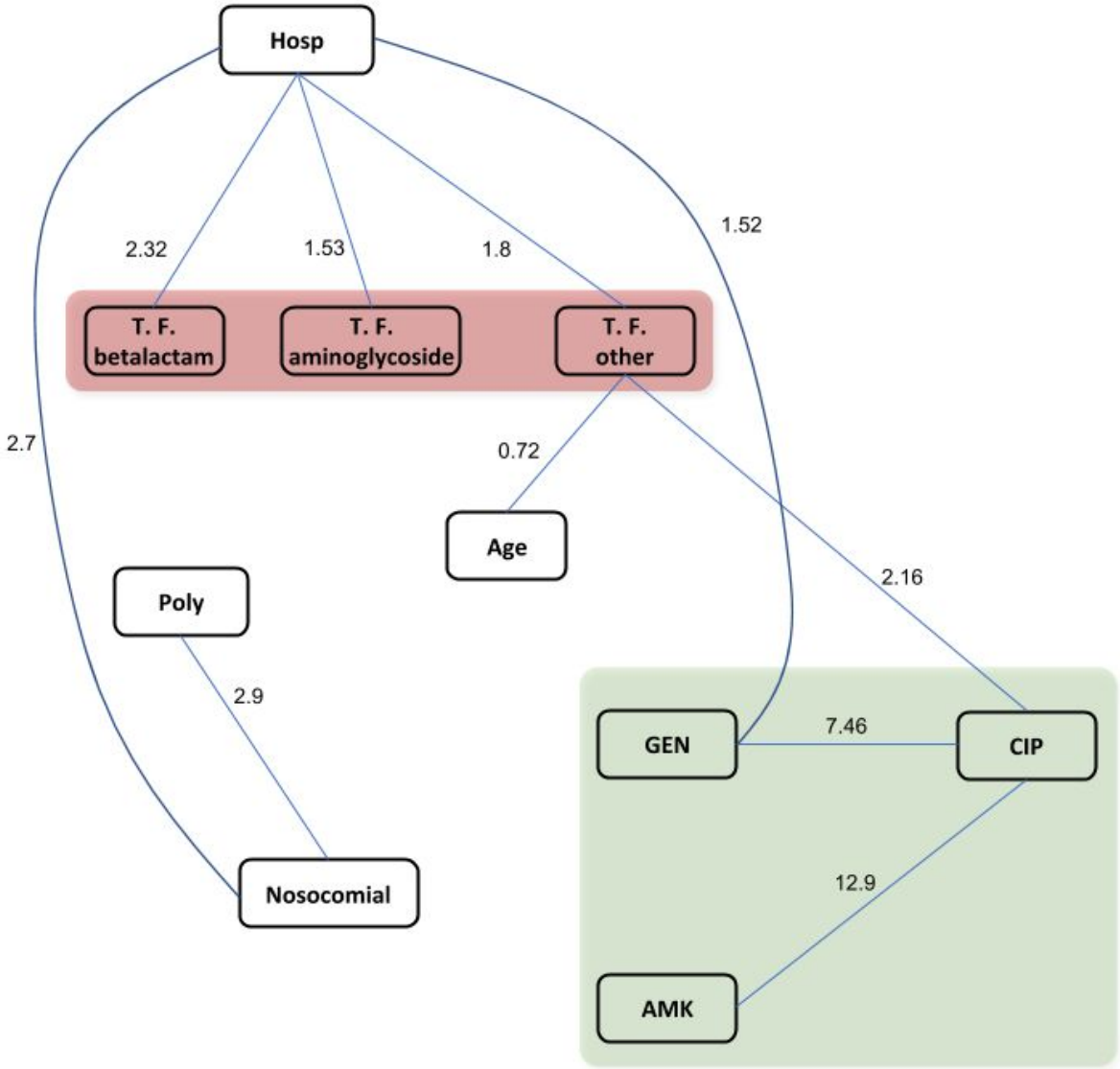


Table 3: *Klebsiella pneumoniae* OR estimates and their 95% credible intervals. None of the credible intervals contains zero.

Variable|Affected by	2.5%	OR	97.5%
T.F.other|Age	0.60	0.72	0.84
Hosp|T.F.aminoglycoside	1.31	1.53	1.78
Hosp|T.F.betalactam	2.00	2.32	2.68
Hosp|T.F.other	1.57	1.80	2.06
Nosocomial|Hosp	2.19	2.70	3.38
Poly|Nosocomial	2.06	2.90	4.09
AMK|CIP	6.06	12.90	35.65
CIP|T.F.other	1.50	2.16	3.09
CIP|GEN	5.10	7.46	11.12
GEN|Hosp	1.28	1.52	1.80

Note: T.F. prefix denotes having taken an antibiotic from the given class; Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; AMK = amikacin; CIP = ciprofloxacin; GEN = gentamicin.

Figure 3: ABN model for *P. aeruginosa*. Arcs denote direct links between variables, and adjacent values are ORs derived from the model (see Table 4). Nodes with antibiotic names denote resistance to those antibiotics (enclosed in a green rectangle). T.F. prefix denotes variables indicating having taken an antibiotic from the given class in the prior year (enclosed in a red rectangle); Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; CAZ = ceftazidime; GEN = gentamicin; IPM = imipenem.

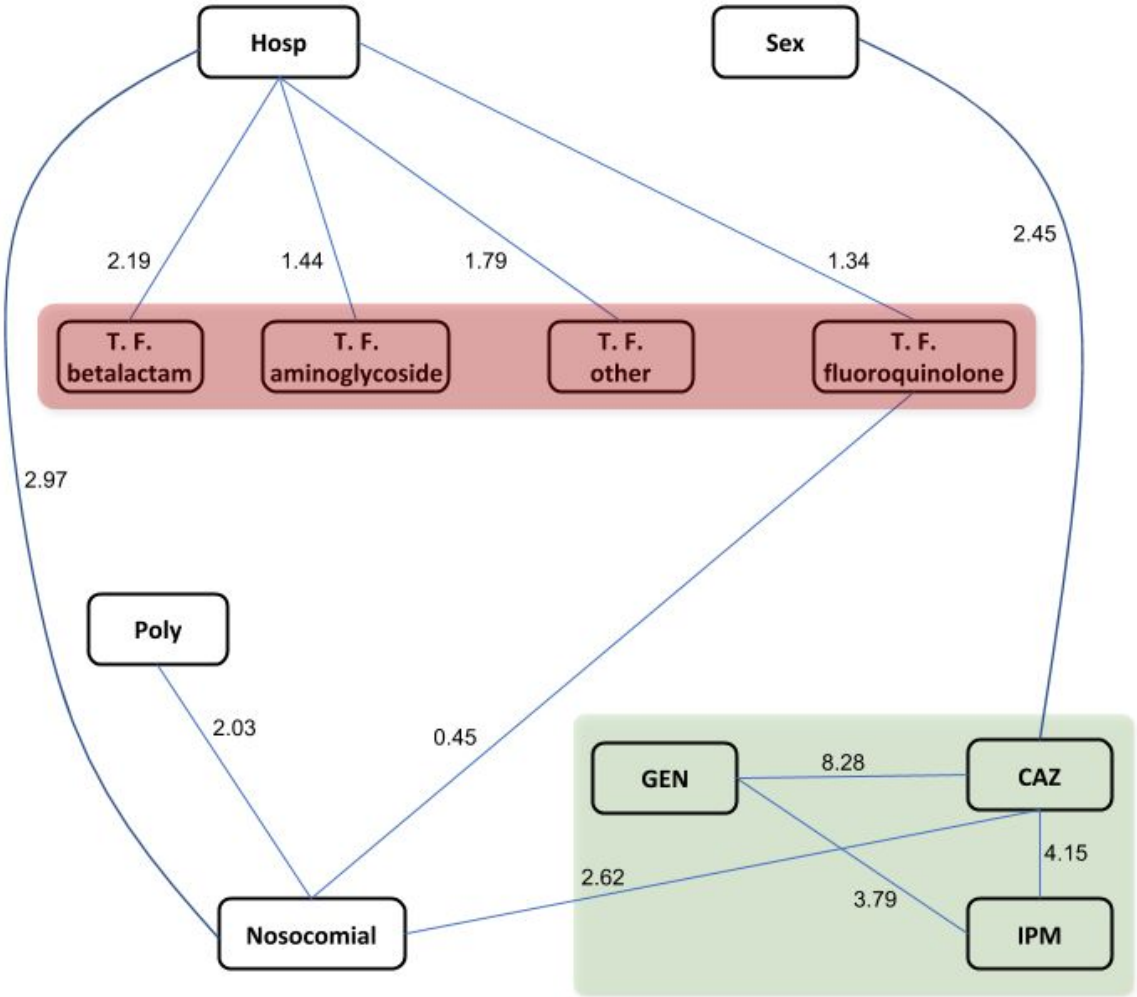


Table 4: *Pseudomonas aeruginosa* OR estimates and their 95% credible intervals. None of the credible intervals contains zero.

Variable|Affected by	2.5%	OR	97.5%
Hosp|T.F.aminoglycoside	1.21	1.44	1.70
Hosp|T.F.fluoroquinolone	1.14	1.34	1.56
Hosp|T.F.betalactam	1.82	2.19	2.61
Hosp|T.F.other	1.52	1.79	2.10
Nosocomial|T.F.fluoroquinolone	0.28	0.45	0.69
Nosocomial|Hosp	2.37	2.97	3.82
Poly|Nosocomial	1.43	2.03	2.88
CAZ|Sex	1.37	2.45	4.74
CAZ|Nosocomial	1.48	2.62	4.88
GEN|CAZ	4.71	8.28	14.46
IPM|CAZ	2.24	4.15	7.51
IPM|GEN	2.09	3.79	6.71

Note: T.F. prefix denotes having taken an antibiotic from the given class; Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; CAZ = ceftazidime; GEN = gentamicin; IPM = imipenem.

Figure 4: ABN model for *P. mirabilis*. Arcs denote direct links between variables, and adjacent values are ORs derived from the model (see Table 5). Nodes with antibiotic names denote resistance to those antibiotics (enclosed in a green rectangle). T.F. prefix denotes variables indicating having taken an antibiotic from the given class in the prior year (enclosed in a red rectangle); Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; CXM = cefuroxime; CIP = ciprofloxacin; GEN = gentamicin; Sulf-Trim = sulfamethoxazole-trimethoprim.

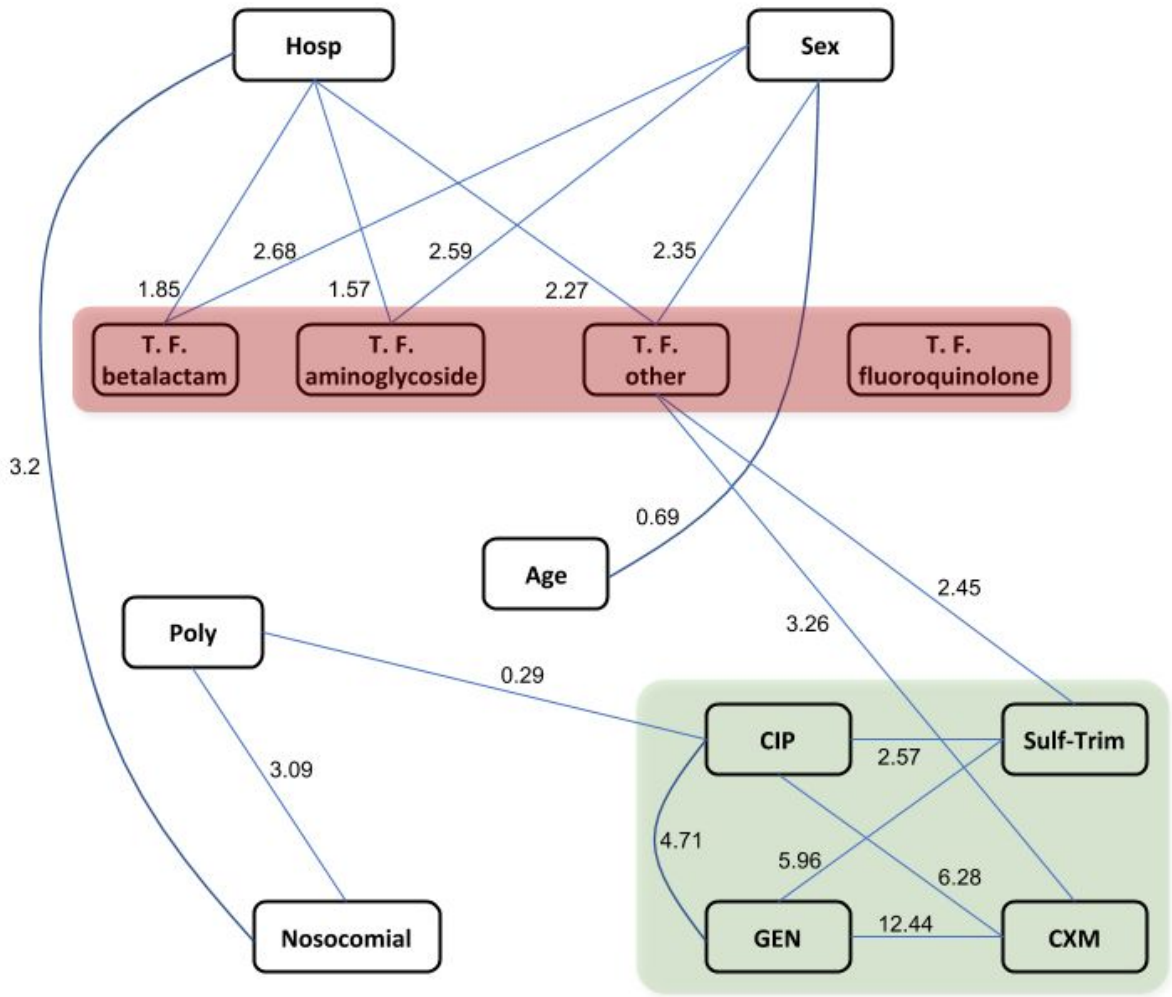


Table 5: *Proteus mirabilis* OR estimates and their 95% credible intervals. None of the credible intervals contains zero.

Variable|Affected by	2.5%	OR	97.5%
Age|Sex	0.55	0.69	0.84
T.F.aminoglycoside|Sex	1.51	2.59	4.68
T.F.betalactam|Sex	1.73	2.68	4.16
T.F.other|Sex	1.51	2.35	3.59
Hosp|T.F.aminoglycoside	1.31	1.57	1.86
Hosp|T.F.betalactam	1.54	1.85	2.20
Hosp|T.F.other	1.91	2.27	2.69
Nosocomial|Hosp	2.46	3.20	4.32
Poly|Nosocomial	2.05	3.09	4.68
CXM|T.F.other	2.16	3.26	4.95
CIP|Poly	0.17	0.29	0.45
CIP|CXM	4.00	6.28	10.19
GEN|CXM	7.47	12.44	21.57
GEN|CIP	2.80	4.71	8.04
Sulf-Trim|T.F.other	1.50	2.45	4.07
Sulf-Trim|CIP	1.51	2.57	4.38
Sulf-Trim|GEN	3.51	5.96	10.58

Note: T.F. prefix denotes having taken an antibiotic from the given class; Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; CXM = cefuroxime; CIP = ciprofloxacin; GEN = gentamicin; Sulf-Trim = sulfamethoxazole-trimethoprim.

Figure 5: Directed acyclic graph of the final reduced model for CoNS. Arcs denote direct links between variables, and adjacent values are ORs derived from the model (see Table 6). Nodes with antibiotic names denote resistance to those antibiotics (enclosed in a green rectangle). T.F. prefix denotes variables indicating having taken an antibiotic from the given class in the prior year (enclosed in a red rectangle); Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; CHL = chloramphenicol; ERY = erythromycin; Fusid = fusidic acid; GEN = gentamicin; OFX = ofloxacin; OXA = oxacillin; Sulf-Trim = sulfamethoxazole-trimethoprim.

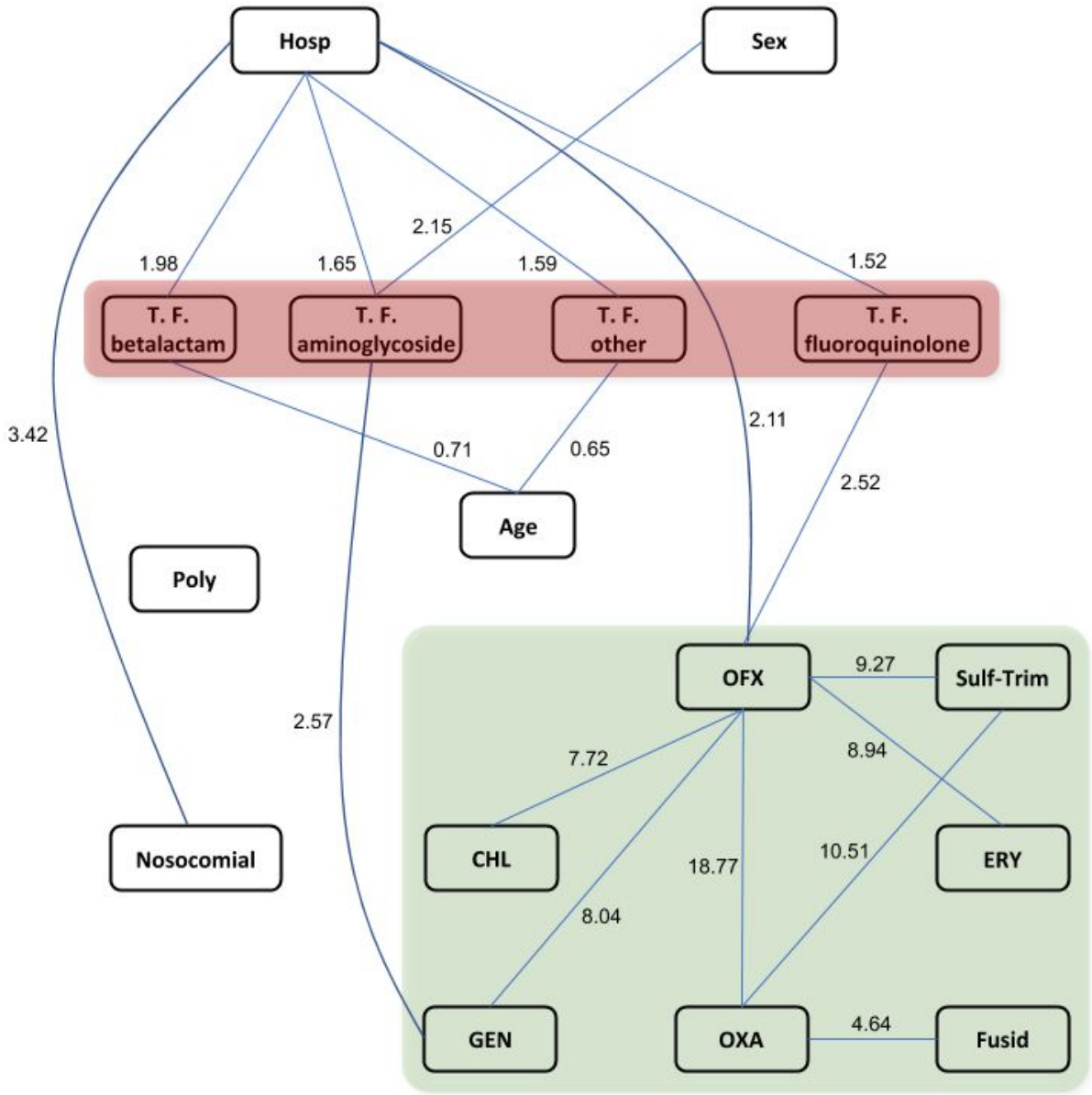
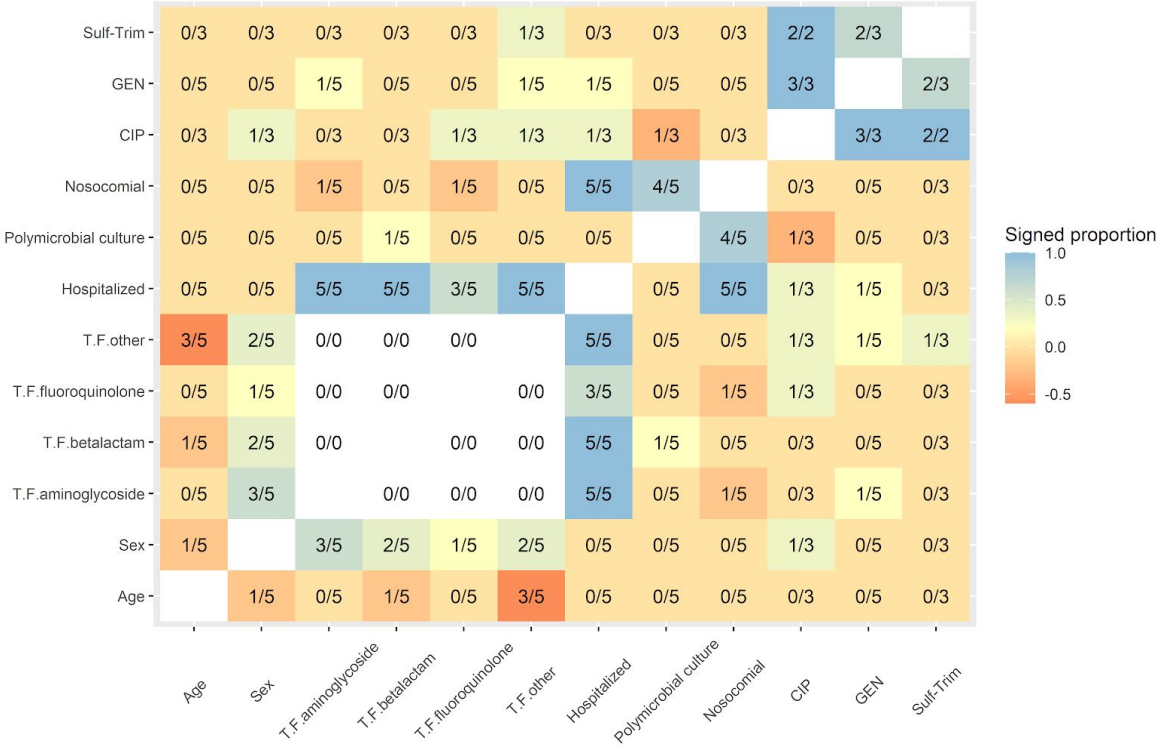


Table 6: *Staphylococcus* coag. neg. group OR estimates and their 95% credible intervals. None of the credible intervals contains zero.

Variable|Affected by	2.5%	OR	97.5%
T.F.aminoglycoside|Sex	1.31	2.15	3.56
T.F.betalactam|Age	0.58	0.71	0.86
T.F.other|Age	0.53	0.65	0.79
Hosp|T.F.aminoglycoside	1.36	1.65	1.98
Hosp|T.F.fluoroquinolone	1.29	1.52	1.78
Hosp|T.F.betalactam	1.67	1.98	2.33
Hosp|T.F.other	1.35	1.59	1.86
Nosocomial|Hosp	2.66	3.42	4.48
CHL|OFX	5.13	7.72	12.03
ERY|OFX	5.80	8.94	14.06
Fusid|OXA	2.88	4.64	7.73
GEN|T.F.aminoglycoside	1.51	2.57	4.42
GEN|OFX	5.09	8.04	13.48
OFX|T.F.fluoroquinolone	1.49	2.52	4.37
OFX|Hosp	1.69	2.11	2.64
OXA|OFX	9.90	18.77	40.75
OXA|Sulf-Trim	5.22	10.51	24.46
Sulf-Trim|OFX	6.11	9.27	14.35

Note: T.F. prefix denotes having taken an antibiotic from the given class; Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; CHL = chloramphenicol; ERY = erythromycin; Fusid = fusidic acid; GEN = gentamicin; OFX = ofloxacin; OXA = oxacillin; Sulf-Trim = sulfamethoxazole-trimethoprim.

Figure 6: Consistency matrix showing proportion of models which contain a given arc. Each cell contains the proportion of models containing both variables and having a direct link between them. Colors correspond to positive or negative links (all signs of links were consistent between the datasets). Variables which appear with only one of the 5 bacterial species were omitted for clarity (but are found in Figure S1). T.F. prefix denotes having taken an antibiotic from the given class; Hosp = log(days hospitalized + 1); Poly = polymicrobial culture; CIP = ciprofloxacin; GEN = gentamicin; Sulf-Trim = sulfamethoxazole-trimethoprim.



In all bacterial species we found a positive, direct link between having taken a betalactam, aminoglycoside, or “other” (i.e., any antibiotic not a member of the three classes which were most frequently used) antibiotic in the prior year, with days hospitalized in the prior year. In addition, three out of five of the analyses found a relationship between days hospitalized in the prior year and having taken a fluoroquinolone antibiotic. These relationships are somewhat expected, as more days spent at the hospital likely positively associated with having received antibiotic treatment, and vice versa.

Age and sex were only directly linked in the *P. mirabilis* dataset (Table 5 and Figure 4), with females being older at the time of the bacterial infection. When analyzing *E. coli* (Table 2 and Figure 1), *K. pneumoniae* (Table 3 and Figure 2), and CoNS (Table 6 and Figure 5), we found that older people were less likely to be prescribed an antibiotic in the “other” category. Older people were also less likely to be prescribed a betalactam in the CoNS model (Table 6 and Figure 5). Males were more likely to be prescribed aminoglycosides (*E. Coli* [Table 2 and Figure 1], *P. mirabilis* [Table 5 and Figure 4], and CoNS [Table 6 and Figure 5]), betalactams and other antibiotics (E. Coli [Table 2 and Figure 1] and *P. mirabilis* [Table 5 and Figure 4]), and fluoroquinolones (*E. Coli* [Table 2 and Figure 1]). Additionally, males had infections with increased resistance to ciprofloxacin in the *E. coli* dataset (Table 2 and Figure

1), and to ceftazidime in the *P. aeruginosa* dataset (Table 4 and Figure 3). Resistance to cefuroxime was also positively linked to age in the *E. coli* dataset.

Having a nosocomial infection was directly linked to the presence of multiple bacteria in the culture in nearly all bacterial species, as has been shown [31,32], with CoNS being the one exception. This is plausible, as CoNS are often isolated as a result of culture contamination [33], and their presence might be associated with a decreased probability of the culture yielding other bacterial species. Correspondingly, the presence of additional bacterial species in a sample containing CoNS was less frequent than in all other bacterial species (7%, see Table 1). Nosocomial infection was also related to the number of days hospitalized during the prior year in all five bacterial species. This is expected, since acquiring a nosocomial infection by definition entails that the patient was hospitalized beforehand. Additionally, nosocomial infections were directly linked to having taken a fluoroquinolone in the *P. aeruginosa* dataset (Table 4 and Figure 3) and having taken an aminoglycoside in the *E. coli* dataset (Table 2 and Figure 1). Polymicrobial cultures were only directly linked to decreased resistance to one drug (ciprofloxacin) in only one bacterial species (*P. mirabilis*; Table 5 and Figure 4*)*.

Discussion

We have performed an additive Bayesian network analysis in an attempt to uncover the underlying dependency structure among patient covariates, measures of antibiotic use and different antibiotic resistance types, in cultures of five bacterial species. Such an analysis has the advantages of not only shedding light on the relationships between all pairs of variables examined, but can identify which variables are directly linked rather than only having indirect associations.

All the connections between resistance to antibiotics identified by our statistical models, in all examined bacterial species, indicated positive associations. That is, resistance to a certain antibiotic was never found to decrease resistance to other antibiotics. Furthermore, some of the associations between resistance to different antibiotics were so strong they could not be included in the ABN models without causing numerical issues. This is expected with very similar antibiotics, but surprisingly appeared between antibiotics of different classes as well (Table S1).

Altogether, our results appear at odds with recent experimental findings. Although cross-resistance is prevalent in experimental settings, even between different classes of antibiotics, cross-sensitivity is also often observed [10–13,34]. Nevertheless, as opposed to experimental data, clinical findings usually identify positive associations between resistance of different antibiotics [35]. However, the reasons for the discrepancy between the experimental and clinical results may be hard to elucidate due to the potential presence of confounders in observational, clinical data. By using ABN modelling, we made a step towards resolving this problem and obtained estimates that reflect innate cross-resistance patterns rather than heavily confounded associations. Hence, our results strengthen the notion that at least some of the cross-resistance patterns observed in clinical settings relate

to innate bacterial mechanisms of resistance acquisition. For instance, the direct cross-resistance we identified between fluoroquinolones and macrolides, between beta-lactams and aminoglycosides, and between aminoglycosides and quinolones, have been previously demonstrated in other clinical settings [36–38].

Our results have practical implications on antibiotic treatment. Knowledge of predictors for antibiotic resistance, and in particular for cross-resistance and previous antibiotic usage, is important for tailoring efficient antibiotic treatment regimes to patients. Combination therapy, for example, where a patient is treated with multiple antibiotics simultaneously, has been shown to yield beneficial results to mitigate resistance [37,39,40]. However, the efficiency of combination therapy can be hampered by acquisition of multiple antibiotic resistance of the infecting bacteria. By having estimates of risks for cross-resistance based on patient demographics and previous antibiotic usage, we can mitigate this risk and seek to treat with combinations of antibiotics which have lower probability of cross-resistance.

Moreover, our results could be used to suggest appropriate drugs for alternating or cycling regimes of antibiotic therapy, which have been hypothesized in various settings as a way to reduce emergence of antibiotic resistance [41–43]. Suggestions for selecting such drugs using estimates inferred from ABNs take into account the clinical characteristics of the patients and circulating bacterial strains. Hence, ABNs can identify patterns of cross-resistance which might be different than those obtained *in vitro* and can provide additional information about the potential efficiency of different antibiotic regimes.

However, the main limitation of our study is that our data lacks previous exposure to antibiotics outside the hospital. Potentially, patients could have been exposed to antibiotic treatment in the community rather than during hospitalization, introducing noise into our estimates of the effects of previous antibiotic usage. However, our previous study on a similar dataset showed that prior antibiotic use in the hospital is a strong predictor for antibiotic resistance which can yield high accuracy predictions even without outpatient antibiotic use [5]. Additionally, the consistency of our estimates between bacterial species, as well as relatively narrow coverage of the estimated CIs indicate that our results should be robust to these mising data.

To conclude, our study provides estimates for cross-resistance between different antibiotics in bacterial pathogens isolated in clinical settings. By using an ABN model, we control for confounders and provide results for direct and indirect links between different patient characteristics and antibiotic resistance. Although our analysis does not provide direct causal effect estimates, it presents a step towards studying causal effects on antibiotic resistance patterns. Revealing the dependency structures of multiple variables simultaneously instead of analysing associations affecting a single dependent variable should provide more robust and less spurious relationships between variables. Hence, our results should be of utility in decision making regarding antibiotic treatments strategies.

Acknowledgements

None to declare.

Funding

None to declare.

Transparency declarations

None to declare.

References

1. Rodríguez-Rojas A, Rodríguez-Beltrán J, Couce A *et al.* Antibiotics and antibiotic resistance: a bitter fight against evolution. *Int J Med Microbiol IJMM* 2013;**303**:293–7.
2. Friedman ND, Temkin E, Carmeli Y. The negative impact of antibiotic resistance. *Clin Microbiol Infect* 2016;**22**:416–22.
3. *Antibiotic Resistance Threats in the United States*. Atlanta, GA, USA: Department of Health and Human Services, CDC, 2019.
4. Yelin I, Snitser O, Novich G *et al.* Personal clinical history predicts antibiotic resistance of urinary tract infections. *Nat Med* 2019;**25**:1143–52.
5. Lewin-Epstein O, Baruch S, Hadany L *et al.* Predicting antibiotic resistance in hospitalized patients by applying machine learning to electronic medical records. *medRxiv* 2020:2020.06.03.20120535.
6. Chatterjee A, Modarai M, Naylor NR *et al.* Quantifying drivers of antibiotic resistance in humans: a systematic review. *Lancet Infect Dis* 2018;**18**:e368–78.
7. CROSS RESISTANCE TO ANTIBIOTICS. *J Am Med Assoc* 1952;**148**:470–1.
8. Obolski U, Dellus-Gur E, Stein GY *et al.* Antibiotic cross-resistance in the lab and resistance co-occurrence in the clinic: Discrepancies and implications in E.coli. *Infect Genet Evol J Mol Epidemiol Evol Genet Infect Dis* 2016;**40**:155–61.
9. Blair JMA, Webber MA, Baylay AJ *et al.* Molecular mechanisms of antibiotic resistance. *Nat Rev Microbiol* 2015;**13**:42–51.
10. Imamovic L, Sommer MOA. Use of Collateral Sensitivity Networks to Design Drug Cycling Protocols That Avoid Resistance Development. *Sci Transl Med* 2013;**5**:204ra132-204ra132.
11. Pál C, Papp B, Lázár V. Collateral sensitivity of antibiotic-resistant microbes. *Trends Microbiol* 2015;**23**:401–7.
12. Lázár V, Nagy I, Spohn R *et al.* Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. *Nat Commun* 2014;**5**:4352.
13. Munck C, Gumpert HK, Wallin AIN *et al.* Prediction of resistance development against drug combinations by collateral responses to component drugs. *Sci Transl Med* 2014;**6**:262ra156-262ra156.
14. Kratzer G, Lewis FI, Comin A *et al.* Additive Bayesian Network Modelling with the R Package abn. *ArXiv Prepr ArXiv191109006* 2019.
15. Ruchti S, Meier AR, Würbel H *et al.* Pododermatitis in group housed rabbit does in

Switzerland—Prevalence, severity and risk factors. *Prev Vet Med* 2018;**158**:114–21.
16. Ludwig A, Berthiaume P, Boerlin P *et al.* Identifying associations in Escherichia coli antimicrobial resistance patterns using additive Bayesian networks. *Prev Vet Med* 2013;**110**:64–75.
17. Hartnack S, Odoch T, Kratzer G *et al.* Additive Bayesian networks for antimicrobial resistance and potential risk factors in non-typhoidal Salmonella isolates from layer hens in Uganda. *BMC Vet Res* 2019;**15**:212.
18. Hidano A, Yamamoto T, Hayama Y *et al.* Unraveling Antimicrobial Resistance Genes and Phenotype Patterns among Enterococcus faecalis Isolated from Retail Chicken Products in Japan. *PLoS ONE* 2015;**10**, DOI: 10.1371/journal.pone.0121189.
19. Waterhouse M, Morton A, Mengersen K *et al.* Role of overcrowding in meticillin-resistant Staphylococcus aureus transmission: Bayesian network analysis for a single public hospital. *J Hosp Infect* 2011;**78**:92–6.
20. Lewis FI, Ward MP. Improving epidemiologic data analyses through multivariate regression modelling. *Emerg Themes Epidemiol* 2013;**10**:4.
21. Kratzer G, Pittavino M, Lewis FI *et al. Abn: An R Package for Modelling Multivariate Data Using Additive Bayesian Networks*., 2019.
22. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2019.
23. Friedman N, Goldszmidt M, Wyner A. Data Analysis with Bayesian Networks: A Bootstrap Approach. *ArXiv13016695 Cs Stat* 2013.
24. Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proc 3rd Int Workshop Distrib Stat Comput* 2003;**124**:1–10.
25. Lewis FI, McCormick BJJ. Revealing the Complexity of Health Determinants in Resource-poor Settings. *Am J Epidemiol* 2012;**176**:1051–9.
26. Poon AFY, Lewis FI, Pond SLK *et al.* Evolutionary Interactions between N-Linked Glycosylation Sites in the HIV-1 Envelope. *PLoS Comput Biol* 2007;**3**, DOI: 10.1371/journal.pcbi.0030011.
27. Poon AFY, Lewis FI, Pond SLK *et al.* An Evolutionary-Network Model Reveals Stratified Interactions in the V3 Loop of the HIV-1 Envelope. *PLoS Comput Biol* 2007;**3**, DOI: 10.1371/journal.pcbi.0030231.
28. Poon AFY, Lewis FI, Frost SDW *et al.* Spidermonkey: rapid detection of co-evolving sites using Bayesian graphical models. *Bioinformatics* 2008;**24**:1949–50.
29. Lycett SJ, Ward MJ, Lewis FI *et al.* Detection of Mammalian Virulence Determinants in Highly Pathogenic Avian Influenza H5N1 Viruses: Multivariate Analysis of Published Data. *J Virol* 2009;**83**:9901–10.
30. Milns I, Beale CM, Smith VA. Revealing ecological networks using Bayesian network inference algorithms. *Ecology* 2010;**91**:1892–9.
31. Siegman-Igra Y, Kulka T, Schwartz D *et al.* Polymicrobial and monomicrobial bacteraemic urinary tract infection. *J Hosp Infect* 1994;**28**:49–56.
32. Weinstein MP, Reller LB, Murphy JR. Clinical importance of polymicrobial bacteremia. *Diagn Microbiol Infect Dis* 1986;**5**:185–96.
33. Altindis M, Koroglu M, Demiray T *et al.* A Multicenter Evaluation of Blood Culture Practices, Contamination Rates, and the Distribution of Causative Bacteria. *Jundishapur J Microbiol* 2016;**9**, DOI: 10.5812/jjm.29766.
34. Adamus-Białek W, Wawszczak M, Arabski M *et al.* Ciprofloxacin, amoxicillin, and aminoglycosides stimulate genetic and phenotypic changes in uropathogenic Escherichia coli strains. *Virulence* 2019;**10**:260–76.
35. Chang H-H, Cohen T, Grad YH *et al.* Origin and Proliferation of Multiple-Drug Resistance in Bacterial Pathogens. *Microbiol Mol Biol Rev* 2015;**79**:101–16.
36. Hwang TJ, Hooper DC. Association between fluoroquinolone resistance and resistance

to other antimicrobial agents among Escherichia coli urinary isolates in the outpatient setting: a national cross-sectional study. *J Antimicrob Chemother* 2014;**69**:1720–2.
37. Monedero I, Caminero JA. Management of multidrug-resistant tuberculosis: an update. *Ther Adv Respir Dis* 2010;**4**:117–27.
38. Tsukamoto N, Ohkoshi Y, Okubo T *et al.* High Prevalence of Cross-Resistance to Aminoglycosides in Fluoroquinolone-Resistant Escherichia coli Clinical Isolates. *Chemotherapy* 2013;**59**:379–84.
39. REX Consortium. Heterogeneity of selection and the evolution of resistance. *Trends Ecol Evol* 2013;**28**:110–8.
40. Vandamme A-M, Camacho RJ, Ceccherini-Silberstein F *et al.* European Recommendations for the Clinical Use of HIV Drug Resistance Testing: 2011 Update. *AIDS Rev* 2011;**13**.
41. Kim S, Lieberman TD, Kishony R. Alternating antibiotic treatments constrain evolutionary paths to multidrug resistance. *Proc Natl Acad Sci U S A* 2014;**111**:14494–9.
42. Obolski U, Hadany L. Implications of stress-induced genetic variation for minimizing multidrug resistance in bacteria. *BMC Med* 2012;**10**:89.
43. Obolski U, Stein GY, Hadany L. Antibiotic Restriction Might Facilitate the Emergence of Multi-drug Resistance. *PLOS Comput Biol* 2015;**11**:e1004340.