


Identifying patient-level risk factors associated with non- β -lactam resistance outcomes in invasive MRSA infections in the United States using chain graphs

William J. Love ^{1*}, C. Annie Wang¹ and Cristina Lanzas¹

¹Department of Population Health and Pathobiology, North Carolina State University, Raleigh, NC, USA

*Corresponding author. E-mail: wjlove@ncsu.edu
 @EpidemiologyDVM

Received 13 December 2021; accepted 2 June 2022

Background: MRSA is one of the most common causes of hospital- and community-acquired infections. MRSA is resistant to many antibiotics, including β -lactam antibiotics, fluoroquinolones, lincosamides, macrolides, aminoglycosides, tetracyclines and chloramphenicol.

Objectives: To identify patient-level characteristics that may be associated with phenotype variations and that may help improve prescribing practice and antimicrobial stewardship.

Methods: Chain graphs for resistance phenotypes were learned from invasive MRSA surveillance data collected by the CDC as part of the Emerging Infections Program to identify patient level risk factors for individual resistance outcomes reported as MIC while accounting for the correlations among the resistance traits. These chain graphs are multilevel probabilistic graphical models (PGMs) that can be used to quantify and visualize the complex associations among multiple resistance outcomes and their explanatory variables.

Results: Some phenotypic resistances had low connectivity to other outcomes or predictors (e.g. tetracycline, vancomycin, doxycycline and rifampicin). Only levofloxacin susceptibility was associated with healthcare-associated infections. Blood culture was the most common predictor of MIC. Patients with positive blood culture had significantly increased MIC of chloramphenicol, erythromycin, gentamicin, lincomycin and mupirocin, and decreased daptomycin and rifampicin MICs. Some regional variations were also observed.

Conclusions: The differences in resistance phenotypes between patients with previous healthcare use or positive blood cultures, or from different states, may be useful to inform first-choice antibiotics to treat clinical MRSA cases. Additionally, we demonstrated multilevel PGMs are useful to quantify and visualize interactions among multiple resistance outcomes and their explanatory variables.

Introduction

Antimicrobial resistance (AMR) poses a significant threat to modern medicine by decreasing the efficacy of antimicrobial treatment and increasing patient adverse outcomes and healthcare costs.^{1,2} Many of the bacterial pathogens of most clinical concern are resistant to multiple classes of antimicrobials and classified as MDR pathogens, among which MRSA is one of the most common causes of hospital-acquired and community-acquired infections. In 2019, the CDC classified MRSA as a serious threat.³ MRSA has become resistant to many antibiotics used for its treatment besides methicillin and most other β -lactam antibiotics, and is frequently resistant to fluoroquinolones, lincosamides, macrolides,

aminoglycosides, tetracyclines and chloramphenicol.⁴ MDR phenotypes often result from accumulating multiple genes encoding for resistance to different single drugs on mobile genetic elements, e.g. plasmids and integrons, and individual genes encoding multi-drug efflux pumps.^{5,6} These genetic elements cause MDR phenotypes to be more common than would be expected and allow for co-selection, which can drive changes in the prevalence of resistance.⁷ For instance, increased prescription of tetracyclines and gentamicin in Europe during the 1980s was associated with an increased prevalence of MRSA.⁸

Clinical infection with MDR pathogens results in reduced therapeutic efficacy of antibiotics and worse patient outcomes.⁹

Accounting for the MDR nature of pathogens is crucial to appropriately identifying risk factors for such infections, and thus improving clinical management. The complex correlation structure among the antimicrobial susceptibility testing (AST) results requires multivariate analysis methods that account for multiple outcomes simultaneously. Standard statistical methods treat each outcome as independent, and most multivariate methods require *a priori* assumptions about outcome correlation structure. Alternative methods that allow for the simultaneous estimation of outcome correlations and causal effect measures are more efficient and promising for MDR outcomes.¹⁰

Probabilistic graphical models (PGMs) are methods for representing complex joint distributions without *a priori* designation of outcome correlation structures.¹¹ In PGMs, nodes represent variables of interest and edges connecting nodes represent an association between those two variables. Bayesian networks are a type of PGM that decompose the joint distribution into a directed acyclic graph (DAG) with directed edges representing conditional associations. In a DAG, each child node is conditionally independent of its non-descendants given its parents (nodes directly connected to the node of interest) (Figure 1a). DAGs are extensively used in epidemiology to represent causal diagrams and identify statistical models most likely to yield unbiased effect estimates.¹² Bayesian networks have been previously applied to evaluate cross-resistance using resistance data from clinical bacterial cultures,¹³ but associations among resistance outcomes are more appropriately represented by undirected edges.¹⁴ Therefore, correlated resistance outcomes are better represented with PGMs by undirected edges, also called Markov networks (Figure 1b). The structure of a Markov network of resistance outcomes can be learned by using partial correlations coefficients followed by a penalization approach to remove the weakest edges.^{14,15}

For evaluating potential risk factors for multiple resistance outcomes, a third type of PGM called chain graphs shows promise. Chain graphs are a hybrid between Bayesian and Markov networks (Figure 1c) and contain both directed and undirected edges. This enables the identification of effect estimates between risk factors and resistances, represented by directed edges, while simultaneously accounting for joint distributions of resistance outcomes, represented by undirected edges. In this study, we used chain graphs to analyse MRSA surveillance data collected by the CDC Emerging Infections Program (CDC-EIP),¹⁶ an active, population-based and laboratory-based surveillance system. The objective of this study was to identify risk factors for individual resistance outcomes while considering the correlations among phenotypic resistant outcomes.

Methods

Data description

The CDC-EIP invasive MRSA dataset contained 12 066 AST records and associated variables from MRSA-infected patients collected from 2005 to 2016. The dataset consisted of AST results for 13 antibiotics (Table 1), and 15 potential predictors describing patients' risk factors (Table 2) for each record. Invasive MRSA cases were collected from eight metropolitan areas in different states and the whole state of Connecticut (Table 3). The MIC values were \log_2 transformed so that a unit change corresponded to a single 2-fold dilution.^{14,17}

No AST results were available for 527 records (4.3%), 2547 additional records (21.1%) were missing chloramphenicol susceptibility results, and an additional 10 records (<0.1%) were missing at least one other AST result. No chloramphenicol susceptibility results were reported for the 1985 isolates collected in 2013 and later. Records missing at least one AST result were omitted from the study, leaving 8982 isolates of MRSA collected between 2005 and 2012. None of the mean \log_2 MIC values in the omitted data significantly differed from the data included by more than one dilution. [Supplementary materials](#) are included to more completely summarize data missingness (Table S1, available as [Supplementary data](#) at JAC-AMR Online) and distribution of AST results over time with tests for linear trend (Figure S1a-m). Both hospital-onset (HO; $n=1927$) and hospital-acquired community-onset (HACO; $n=5201$) isolates were classified as healthcare-associated (HA) isolates; community-associated (CA; $n=1854$) isolates were the least frequent class.

Chain graph models

The chain graph consists of two types of nodes: the set of outcome variables (Y), and the set of predictor variables (X) (Figure 2). Here, X and Y consisted of the 15 patient-level predictors and the 13 \log_2 MIC variables, respectively. The undirected edges among the \log_2 MIC variables are defined by partial correlations (Ω) and the directed edges from the patient-level risk factors to the \log_2 MICs are defined by estimated effects of predictors on the outcomes (β).

Learning chain graphs from data

Analysing large datasets with multiple correlated outcomes presents a particular set of challenges. Large datasets can create challenges for model selection because small effects will appear to be significant. For example, a sample size of 10 000 samples with a standard deviation of \log_2 MIC 3 dilutions used to make 100 comparisons will have over 95% power to identify a \log_2 MIC change of 0.16 dilutions but this change is smaller than would be considered biologically important. Penalization methods have been used in machine learning applications to avoid overfitting.¹⁸ Weak effect estimates are reduced to zero while stronger effects are biased towards zero. Higher penalty values induce in sparser models. Adjusting the penalty allows researchers to tune analysis in an objective manner analogous to adjusting the significance level α . The least absolute shrinkage and selection operator (lasso) and the graphical lasso are commonly used penalization methods.^{18,19}

The data were fitted to chain graph models using the penalized maximum likelihood estimation method.²⁰ The directed edges that compose β are estimated using the debiased lasso with the penalty λ (Figure 2). The undirected edges that compose Ω are estimated using the graphical lasso estimated with the penalty ρ .

This algorithm has the following steps: first, all terms in β representing possible conditional relationships are screened by regressing each outcome on the full set of predictors using the debiased lasso.²¹ Second, the algorithm finds initial estimates for the β terms, then applies the graphical lasso to the residuals to estimate the partial correlations that comprise Ω . Third, the algorithm returns to estimate β conditioned on the newly estimated Ω , and then again estimates Ω from the residuals. The algorithm iterates through the second and third steps, estimating β conditioned on Ω then Ω from the residuals, until convergence.

The stability approach to regularization selection (StARS) was used to identify penalty values λ and ρ that produced the most stable set of edges.²² Penalties between 0.10 and 0.40 in increments 0.01 were tested for λ and ρ each. The models with penalties below 0.10 were too dense to be easily interpretable, and penalties above 0.40 produced models were too sparse to be informative. Twenty subsamples of 80% of the usable data were taken for each combination. Standard errors for the effect estimates for the selected model were estimated using residual mean square errors.

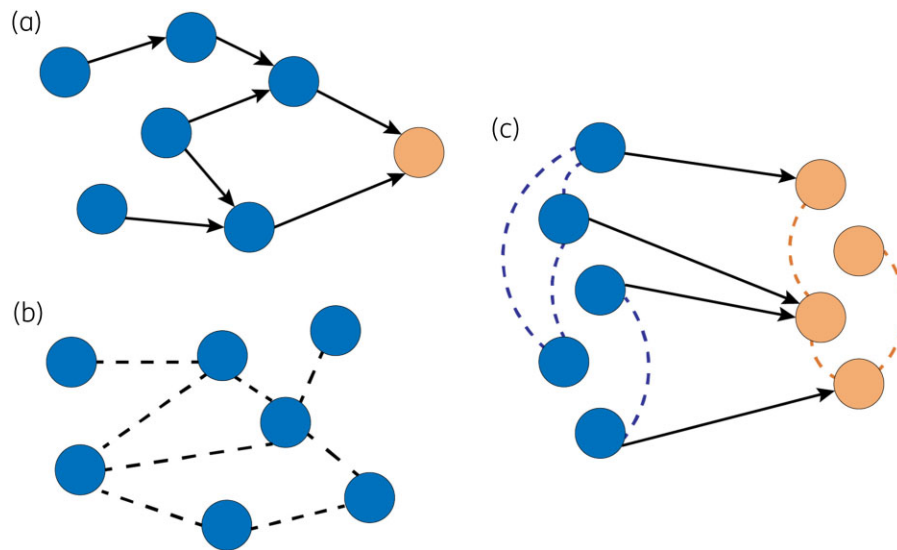


Figure 1. Comparison of three types of probabilistic graphical models. (a) Bayesian networks have directed edges connecting nodes and are commonly used in epidemiology in the form of directed acyclic graphs, which illustrate the relationship between exposure variables (blue nodes) and outcome variables (orange node), as shown here. (b) Markov networks have undirected edges connecting nodes within a graph, here depicted with a dotted line. (c) Chain graphs utilize both directed and undirected edges, with undirected edges connecting nodes within the same layer (denoted by the respective orange and blue colour schemes) and directed edges connecting nodes between different layers.

Analyses were performed using R (v4.1.3).²³ Chain graph learning code was adapted from Lin *et al.* (2016).^{20,24}

Results

The most common non-susceptible phenotypes were CHL (40.4%), CLI (42.5%), ERY (91.8%) and LEV (81.6%) (Table 1). No other phenotypic resistances exceeded 10% prevalence. No isolates were vancomycin resistant, and three isolates had intermediate susceptibility to vancomycin (MIC_{VAN} = 4 mg/L).

Stability selection found the most stable model structure occurred with penalties $\lambda = 0.20$ and $\rho = 0.30$, which produced a model

with 13 edges in the undirected Ω component (16.7% density) and 54 edges in the directed β component (18.1% density) (Figure 3). The model was divided into neighborhood graphs to facilitate visualization (Figure 4). Each outcome node is displayed with its adjacent nodes. Overall, the total number of adjacent nodes was highly variable across resistance outcomes; VAN and TET had only 1 edge, while CHL had the most with 14 edges.

The directed edges in the graphical model represent the debiased non-zero estimated linear coefficients describing the average effect measure of the patient predictors on the outcome MICs (Table 4). Eight of the 15 predictors had no significant relationship with any of the susceptibilities. Eleven of

Table 1. Proportions of non-susceptible isolates and breakpoints for 13 phenotypic susceptibilities in 8982 MRSA isolates collected in the USA between 2005 and 2012

Drug	Code	Mean log ₂ MIC (±SD)	Breakpoint (mg/L)	Log ₂ breakpoint	Non-susceptible (%)
Chloramphenicol	CHL	3.40 ± 0.51	≥16	≥4	40.4
Clindamycin	CLI	0.58 ± 3.01	≥1	≥0	42.5
Erythromycin	ERY	2.67 ± 1.16	≥1	≥0	91.8
Daptomycin	DAP	-0.95 ± 0.28	≥2	≥1	0.5
Doxycycline	DOX	0.01 ± 0.49	≥8	≥3	1.6
Tetracycline	TET	0.14 ± 0.88	≥8	≥3	4.3
Gentamicin	GEN	1.06 ± 0.64	≥8	≥3	2.6
Levofloxacin	LEV	2.73 ± 1.91	≥2	≥1	81.6
Linezolid	LIN	1.30 ± 0.48	≥8	≥3	<0.1
Mupirocin	MUP	2.17 ± 0.93	≥8	≥3	4.9
Rifampicin	RIF	-0.92 ± 0.57	≥2	≥1	2.6
Trimethoprim/sulfamethoxazole	TMS	-0.72 ± 1.06	≥4	≥2	7.6
Vancomycin	VAN	-0.19 ± 0.47	≥4	≥2	0.0

Breakpoints are based on CLSI published MIC breakpoints.⁴³

Table 2. Summary of patient covariates for 8982 MRSA isolates collected in the USA between 2005 and 2012

Predictor/Level	Data code	n	(%)
MRSA-positive blood culture	BLOOD	7442	83%
MRSA-associated disease			
Abscess	ABSCESS	502	6%
Bursitis	BURS	187	2%
Cellulitis	CELL	846	9%
Urinary tract infection	UTI	580	6%
Internal surgical site	SURGSITE	182	2%
Sepsis	SEPTIC	606	7%
Chronic dialysis at culture	CIDAL9	654	7%
MRSA infection type ^a			
CA	class_x	1854	21%
HO		1927	21%
HACO		5201	58%
Conditions in the previous year			
MRSA culture	PREVMRSA	2276	25%
Hospitalization	HOSPYR	5312	59%
Long-term care facility	LTCYR	2265	25%
Dialysis	DIALYR	1522	17%
Location			
California	CA	1290	14%
Colorado	CO	540	6%
Connecticut	CT	711	8%
Georgia	GA	2013	22%
Maryland	MD	429	5%
Minnesota	MN	1147	13%
New York	NY	1276	14%
Oregon	OR	1130	13%
Tennessee	TN	446	5%
Year			
2005–08	Year	4964	55%
2009–12		4018	45%

^aCA, community-associated; HO, hospital-onset; HACO, hospital-associated community onset.

13 susceptibilities were affected by at least one predictor, and 9 of those susceptibilities had at least one regional predictor (Table 4). The undirected edges in the graphical model represent the penalized partial correlations between the outcome MICs (Table 5).

The model found two susceptibilities (TET and DOX) without any significantly associated covariates (Figure 4 and Table 4). Patients with positive blood cultures had significantly increased MICs ($\alpha=0.05$) for five tested susceptibilities ($\beta_{\text{CHL|BLOOD}}=2.72$ dilutions, $\beta_{\text{ERY|BLOOD}}=1.86$ dilutions, $\beta_{\text{GEN|BLOOD}}=0.97$ dilutions, $\beta_{\text{LIN|BLOOD}}=1.19$ dilutions, $\beta_{\text{MUP|BLOOD}}=1.88$ dilutions) and significantly decreased MICs in two ($\beta_{\text{DAP|BLOOD}}=-0.90$ dilutions, $\beta_{\text{RIF|BLOOD}}=-0.87$ dilutions). On average, samples from 2009 and later had significantly lower VAN ($\beta_{\text{VAN|YR}>2009}=-0.42$ dilutions). All tested susceptibilities except VAN and DAP were associated with at least one state. Estimated standard errors varied from 0.01 to 0.03 dilutions with $P \ll 0.001$ in all cases.

Discussion

The chain graph model fit to the surveillance data showed complex and varied relationships between patient predictors and measured susceptibilities. The neighborhoods for TMS and VAN were simple, being only affected by a single predictor each (region and time, respectively), and were independent of all other susceptibilities (Figure 4). Other susceptibilities, such as CHL, had much more complex associations with multiple predictors and were correlated with many other resistances (Figure 4). It appears likely that within the sampled population of MRSA isolates, different types of drivers are responsible for the resistance phenotypes noted. The chain graphs were learned with phenotypic data and without any genotypic or genomic information. Hence, any interpretations we provide as to the genetic and evolutionary mechanisms responsible for noted are limited to speculation that is as consistent with our current understand of MRSA population genetics as possible.

IDSA guidelines recommend clindamycin, combination folate pathway inhibitors and doxycycline to treat uncomplicated skin and soft tissue MRSA infections.²⁵ DOX was associated with TET ($\omega_{\text{DOX-TET}}=0.43$) and weakly with GEN ($\omega_{\text{DOX-GEN}}=0.01$). DOX and TET had the highest partial correlation among resistances, which would be expected given the structural and functional similarity between both drugs. A possible, though still speculative, explanation for the association between gentamicin and doxycycline both bind to the 30S ribosomal subunit, so some mutations may affect both resistances simultaneously.

Trimethoprim/sulfamethoxazole showed only regional variation, with all eight states other than GA having average TMS just under one dilution less than the isolates from GA. Isolates from patients with MRSA-associated cellulitis had slightly lower CLI on average ($\beta_{\text{CLI|Cellulitis}}=-0.76$ dilutions) and showed some regional variation with four states. The only resistance associated with CLI was LEV ($\omega_{\text{CLI-LEV}}=0.05$), which is relatively weak and co-selection is unlikely to be a meaningful driver of clindamycin resistance.

Vancomycin is commonly recommended as treatment for a variety of MRSA infections including bacteraemia, endocarditis, pneumonia or bone and joint infections.²⁵ Only year was a meaningful predictor of VAN, indicating a slight decrease in mean VAN in isolates during or after 2009 compared with those collected prior to 2009 ($\beta_{\text{VAN|YR}\geq 2009}=-0.42$ dilutions); nearly all this variation occurred within the susceptible MIC range ($\text{VAN} \leq 2$ mg/L). Other susceptibilities were not associated with VAN. Despite prior reports that *in vitro* exposure to vancomycin may select for higher DAP,²⁶ penalization in the current study reduced $\omega_{\text{DAP-VAN}}$ to zero and no isolates were concurrently daptomycin and vancomycin non-susceptible.

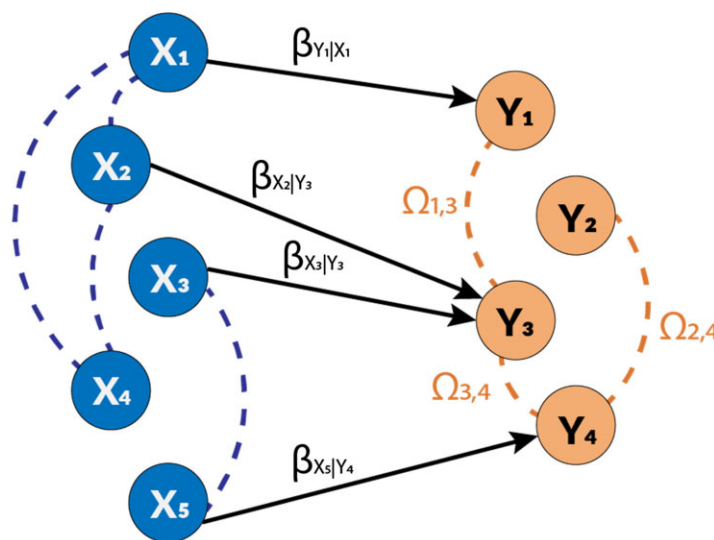
Other drugs used to treat invasive MRSA cases include linezolid, rifampicin and daptomycin.²⁵ The penalized partial correlations between these MICs and other resistances were small, except with CHL, which is discussed below. Very weak negative partial correlations were found between DAP and LIN ($\omega_{\text{DAP-LIN}}=-0.07$) and MUP ($\omega_{\text{DAP-MUP}}=-0.02$). A very weak positive partial correlation with RIF ($\omega_{\text{DAP-RIF}}<0.01$) was noted, meaning that rifampicin and daptomycin use are unlikely to meaningfully impact resistances to each other.

Table 3. Distribution of 8982 invasive MRSA samples from the 9 CDC-EIP sampling sites

State	Year								Total
	2005	2006	2007	2008	2009	2010	2011	2012	
CA	342	250	148	140	151	122	120	17	1290
CO	78	3	62	58	105	115	119	0	540
CT	142	97	100	101	94	100	77	0	711
GA	265	178	184	345	345	284	320	92	2013
MD	0	0	0	16	125	152	136	0	429
MN	68	84	76	281	206	191	241	0	1147
NY	283	142	272	280	95	98	90	16	1276
OR	191	147	143	179	167	130	173	0	1130
TN	104	78	82	45	50	51	30	6	446
Total	1473	979	1067	1445	1338	1243	1306	131	8982

The negative correlations between CHL and RIF and DAP are likely attributable to difference in strains, where determinants of chloramphenicol susceptibility do not coincide with determinants of rifampicin susceptibility, e.g. *rpoB*²⁷ or determinants of

daptomycin susceptibility, e.g. *dltABCD*, *mprF* and others.²⁸ Mupirocin resistance can be conferred via the genes *mupA* and *mupB* and via *ileS* mutations.^{29,30} These mechanisms do not appear to provide cross-resistance to other antimicrobial classes,



Symbol	Notation	Definition
	Y	set of outcome variables
	X	set of predictor variables
	β	estimand causal effects of predictor on outcomes
	Ω	partial correlations among outcome variables

Figure 2. Symbolology for vertices and directed and undirected components of chain graphs.

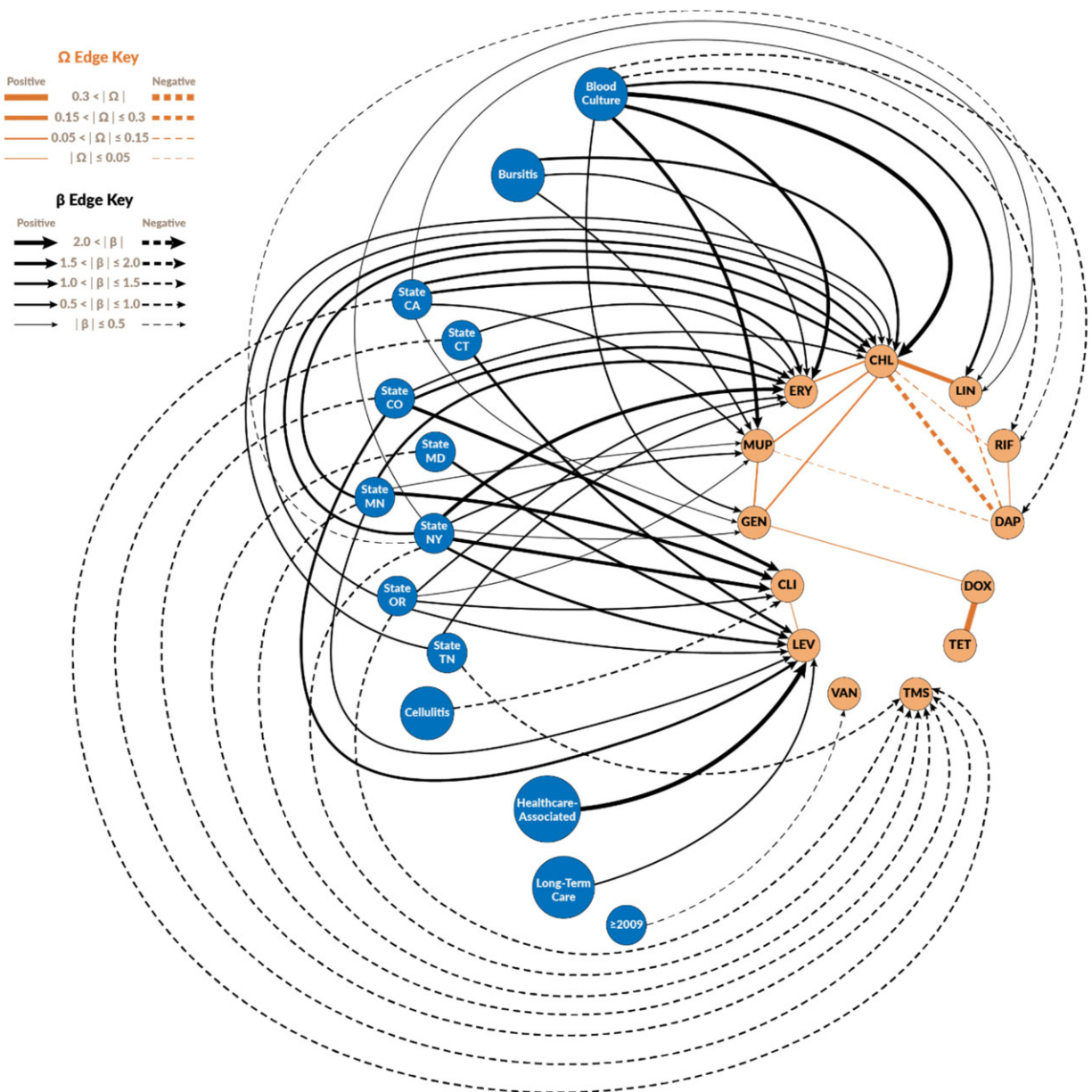


Figure 3. Chain graph representation of the CDC-EIP invasive MRSA surveillance data. Orange nodes represent the resistance outcomes (as \log_2 MIC), and blue nodes are the risk factors associated with the \log_2 MIC nodes. The edges are decorated with line weights and styles to indicate the magnitude of the partial correlations among resistances (Ω) and the magnitude of the effect estimates (β). The vertices representing states (CA, CO, CT, MD, MN, NY, OR and TN) represent the comparison of these states to GA, the referent category, and the node ‘ ≥ 2009 ’ represents the comparison of isolates collected from 2009–12 to isolates collected from 2005–08.

but *mupA* and *mupB* can reside on plasmids with other susceptibility determinants.^{31,32} However, no references describing plasmids carrying both chloramphenicol and mupirocin resistance genes could be found. One possible explanation for the remainder of the CHL neighborhood could be multiresistance plasmids similar to pSCFS1,³³ which carries *lsa(B)* that induces low-level linezolid resistance and *erm(33)* that induces resistance to MLS antibiotics. In this case the plasmid could harbour another gene that induces low-to-moderate chloramphenicol resistance,

e.g. *cat* or *fex*, instead of the *cfr* gene identified on pSCFS1. This combination of genes on a resistance plasmid would explain the moderate to strong positive partial correlations observed between CHL and LIN (0.26) and CHL and ERY (0.14), all of which target the 50S ribosomal subunit. The difficulty in identifying the mechanisms underlying correlations with CHL in this population of MRSA demonstrates the limitations of using phenotypic data alone to interpret noted patterns and the need for more work incorporating genetic data into these chain graphs.

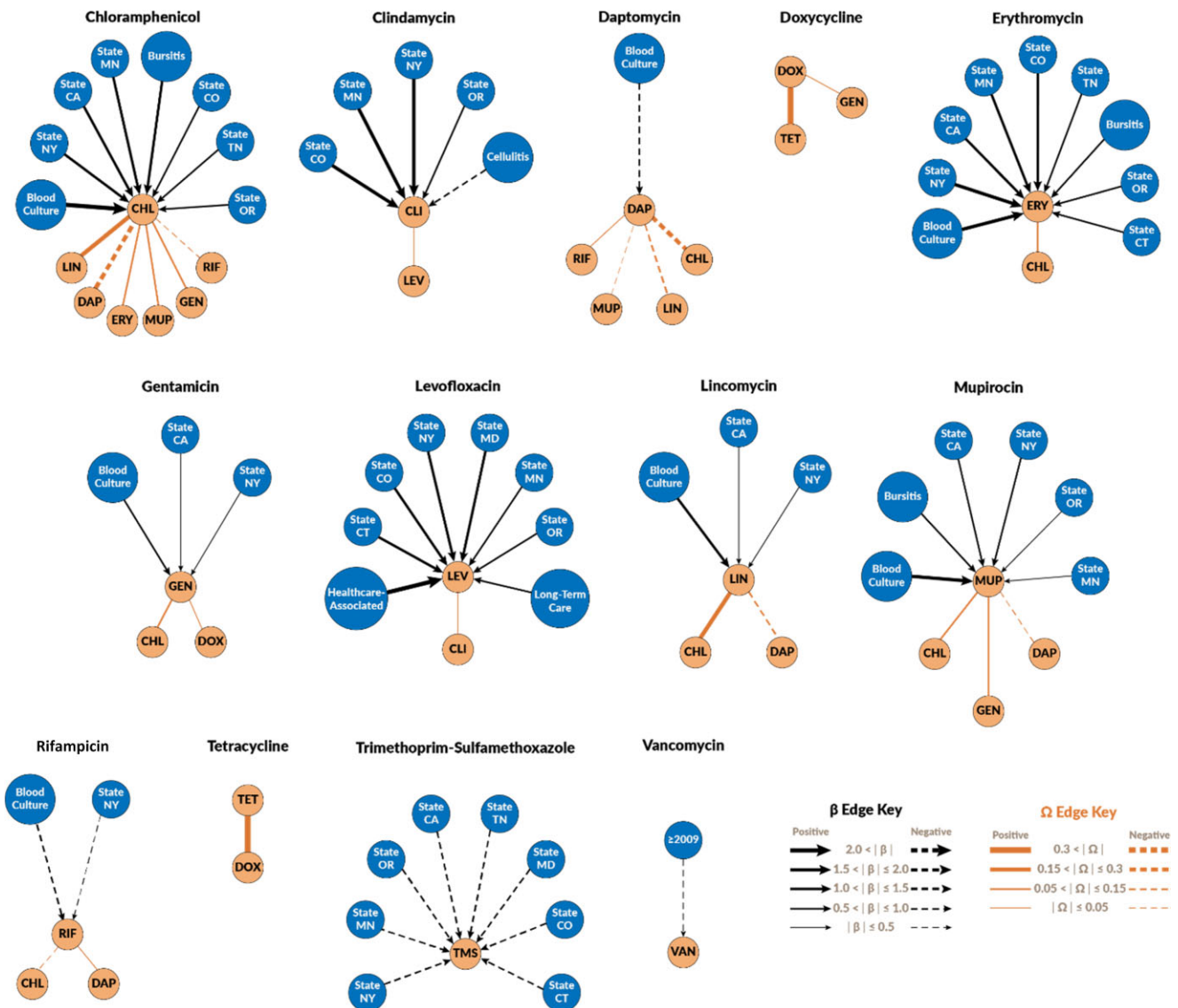


Figure 4. Neighborhood graphs for the chain graph representing the CDC-EIP invasive MRSA surveillance data. Orange nodes represent the resistance outcomes (as \log_2 MIC), and blue nodes are the risk factors associated with the \log_2 MIC nodes. The edges are decorated with line weights and styles to indicate the magnitude of the partial correlations among resistances (Ω) and the magnitude of the effect estimates (β). The vertices representing states (CA, CO, CT, MD, MN, NY, OR and TN) represent the comparison of these states to GA, the referent category, and the node ‘ ≥ 2009 ’ represents the comparison of isolates collected from 2009–12 to isolates collected from 2005–08.

In this surveillance data, 36.8% of the CA-MRSA isolates and 13.9% of HA-MRSA isolates were susceptible to levofloxacin. HA isolates had average LEV values more than two dilutions higher than community associated infections ($\beta_{LEV|hoSp} = 2.20$). Patients with a history of hospitalization in the previous year had a 0.56 dilution increase in LEV. These findings are consistent with previous studies of MDR MRSA in which fluoroquinolone resistance was common in HA-MRSA clones³⁴ and resistance rapidly evolved when fluoroquinolones were used to treat MRSA.³⁵

The MRSA clonal groups found in healthcare-associated infections are generally thought to be distinct from those responsible for community-acquired infections.³⁶ However, LEV was the

only phenotypic susceptibility found to be different between healthcare- and community-associated isolates. HA-MRSA strains tend to be more multidrug resistant because the integrons found in HA-MRSA often contain several gene cassettes with resistance genes conferring resistance to different drug classes.³⁶ Three plausible explanations for why hospital-associated versus community-associated was not identified as a risk factor in our analysis except for LEV are (i) that the epidemiological definition of HA-MRSA and CA-MRSA resulted in misclassified strains; (ii) that the CA-MRSA strains are also likely to contain additional resistances acquired through plasmids; or (iii) that similar strains are circulating in both healthcare settings and the community.³⁶ For instance, USA300 is a strain originally linked to

Table 4. Directed effect estimates (β) of patient factors (X) on \log_2 susceptibilities (Y) from the chain graph model learned from 8962 MRSA isolates

Y	X	β	Standard error
CHL	MRSA-positive blood culture	2.72	0.03
CHL	MRSA-associated bursitis	1.02	0.08
CHL	California	1.22	0.04
CHL	Colorado	0.95	0.05
CHL	Minnesota	1.05	0.04
CHL	New York	1.42	0.04
CHL	Oregon	0.68	0.04
CHL	Tennessee	0.71	0.05
CLI	MRSA-associated cellulitis	-0.73	0.10
CLI	Colorado	1.88	0.13
CLI	Minnesota	1.65	0.09
CLI	New York	1.57	0.09
CLI	Oregon	0.75	0.09
ERY	MRSA-positive blood culture	1.86	0.04
ERY	MRSA-associated bursitis	0.95	0.10
ERY	California	1.34	0.05
ERY	Colorado	1.12	0.06
ERY	Connecticut	0.82	0.06
ERY	Minnesota	1.12	0.05
ERY	New York	1.51	0.05
ERY	Oregon	0.94	0.05
ERY	Tennessee	0.96	0.07
DAP	MRSA-positive blood culture	-0.90	0.01
GEN	MRSA-positive blood culture	0.97	0.02
GEN	California	0.40	0.02
GEN	New York	0.33	0.02
LEV	Long-term care residence in the previous year	0.56	0.05
LEV	HO/HACO MRSA	2.20	0.05
LEV	Colorado	1.33	0.08
LEV	Connecticut	1.35	0.07
LEV	Maryland	1.15	0.09
LEV	Minnesota	0.97	0.06
LEV	New York	1.21	0.06
LEV	Oregon	0.73	0.06
LIN	MRSA-positive blood culture	1.19	0.02
LIN	California	0.28	0.02
LIN	New York	0.41	0.02
MUP	MRSA-positive blood culture	1.88	0.03
MUP	MRSA-associated bursitis	0.89	0.09
MUP	California	0.74	0.04
MUP	Minnesota	0.42	0.04
MUP	New York	0.69	0.04
MUP	Oregon	0.48	0.04
RIF	MRSA-positive blood culture	-0.87	0.02
RIF	New York	-0.33	0.02
TMS	California	-0.91	0.04
TMS	Colorado	-0.84	0.05
TMS	Connecticut	-0.82	0.04
TMS	Maryland	-0.89	0.05

Continued

Table 4. Continued

Y	X	β	Standard error
TMS	Minnesota	-0.96	0.04
TMS	New York	-0.99	0.04
TMS	Oregon	-0.95	0.04
TMS	Tennessee	-0.90	0.05
VAN	≥ 2009	-0.42	0.01

CA-MRSA but has recently been identified as a source of HA-MRSA.³⁷

Patients with positive blood cultures had significantly higher average CHL, ERY, GEN, LIN and MUP values and lower average DAP and RIF. Bacteraemia is a significant source of morbidity and mortality in MRSA infections.^{2,3} Standard of care for treating MRSA bacteraemia is daptomycin and vancomycin.²⁵ Here, DAP and RIF were lower in patients with positive blood cultures, and not significantly different in septic patients. Also, resistance to vancomycin was rare and not associated with blood culture or sepsis. These findings reinforce the recommended practices remain sound for treating these cases.

We expect these chain graphs have two primary applications. First, these chain graphs can be used as surveillance tools. Antimicrobial resistance surveillance typically evaluates how phenotypic susceptibilities change over time, but rarely reports relationships between the phenotypes. Monitoring correlations in phenotypic resistance, especially novel correlations, may identify selection pressures and strain emergence with clinically relevant resistance outcomes. Our previous work has recommended resistance relationship networks, or Rnets, to monitor correlations among phenotypic resistances using widely available AST data.^{14,15} These chain graphs of AMR extend and improve up Rnets by including patient- and isolate-level covariates in the model. The current study serves as a proof-of-concept for facilitating surveillance efforts.

Table 5. Penalized partial correlation estimates (ω) from the chain graph model learned from 8962 MRSA isolates

Susceptibilities	ω	
CHL	ERY	0.14
CHL	DAP	-0.19
CHL	GEN	0.06
CHL	LIN	0.26
CHL	MUP	0.14
CHL	RIF	-0.04
CLI	LEV	0.05
DAP	LIN	-0.07
DAP	MUP	-0.02
DAP	RIF	<0.01
DOX	TET	0.43
DOX	GEN	0.01
GEN	MUP	0.11

The second application is to supplement antibiograms. While these tables are useful to summarize large amounts of clinical data for practitioners, they give no information about the relationships amongst the susceptibilities, which we believe has unrealized potential for clinical decision-making. The relationships between the patient-level covariates and AST results in the β component can also help empirical antimicrobial selection. For example, a clinician presented with a patient with presumptive invasive MRSA and a history of gentamicin administration may be less inclined to prescribe mupirocin prior to seeing AST results since these two susceptibility traits have a positive partial correlation. This clinician may be even less likely to prescribe mupirocin in this patient if they were in California, Minnesota, New York or Oregon, or the patient had a positive blood culture, all of which were found to be factors associated with increased MUP values. These chain graphs provide an accessible and efficient modality for communicating this information to support clinical decision-making. The method can also be extended to include resistance genotypes and virulence factors in addition to AST results, though the necessary genomic data were not available in the current dataset.

There are several limitations in this study. The partial correlations among phenotypic resistance that remain after accounting for the patient-level risk factors can be due to multiple factors. Prescription patterns, community patterns of comorbid illness and genetic structure of MRSA isolates can all influence how drug resistances interact.^{38,39} Without additional information, especially genomic data, these factors cannot be elucidated in the current study. Another limitation of the study is the generalizability of data and the learned structure of the derived PGM to the broader population of MRSA isolated from US patients. The isolates in the current study were collected by the CDC-EIP, an active, population-based surveillance system in nine metropolitan areas.^{16,40} In 2019, the study's referent population was estimated to be approximately 16 million people, or 5% of the US population.¹⁶ Previous studies using this data have suggest the hospital-focused approach may under-represent CA infections.^{40,41} This surveillance programme is restricted to invasive infections and excludes MRSA isolates from skin and soft tissue infections.⁴⁰

There have been a few previous studies that have approached AMR as a multivariate problem, but they have not applied a chain graph or comparable approach. One previous study applied principal component analysis and factor analysis to identify groups of antibiotics with similar trends in MIC from 15 groups of microbes from 17 studies.¹⁷ Rotated principal components correspond to dense subregions in PGMs.¹⁵ The PGMs give a more detailed picture of how the variables are correlated in a way that is visually interpretable and more detailed and that requires fewer assumptions. Bayesian networks have been previously applied to identify risk factors for multiple resistance outcomes but are unsuitable.^{13,42} Chain graphs are a new analytical approach that can provide insights on the interplay between multiple resistance outcomes and their risk factors.

Acknowledgements

We thank Runa H. Gokhale, Amy Gargis, Isaac See and Joseph Lutgring from the CDC for providing useful comments and Jiahe Lin and George Michailidis for providing critical R code.

Funding

This work was supported by the US National Institutes of Health (NIH) (R35GM134934 and F30OD030022) and the Centers for Disease Control and Prevention (CDC) (U01CK000587). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Transparency declarations

None to declare.

Supplementary data

Table S1 and Figure S1 are available as [Supplementary data](#) at *JAC-AMR* Online.

References

- 1 Canton R, Ruiz-Garbajosa P. Co-resistance: an opportunity for the bacteria and resistance genes. *Curr Opin in Pharm* 2011; **11**: 477–85.
- 2 WHO. Antimicrobial Resistance: Global Report on Surveillance. 2014. <https://apps.who.int/iris/rest/bitstreams/515657/retrieve>.
- 3 CDC. Antibiotic Resistance Threats in the United States. 2019. <http://www.cdc.gov/drugresistance/Biggest-Threats.html>.
- 4 Vestegard M, Frees D, Igmer H. Antibiotic resistance and the MRSA problem. *Microb Spectrum* 2019; **7**: 7.2.18.
- 5 Nikaido H. Multidrug resistance in bacteria. *Ann Rev Biochem* 2009; **78**: 119–46.
- 6 Alekshun MN, Levy SB. Molecular mechanisms of antibacterial multidrug resistance. *Cell* 2007; **128**: 1037–50.
- 7 Lehtinen S, Blankquart F, Lipsitch M *et al*. On the evolutionary ecology of multidrug resistance in bacteria. *PLoS Pathogens* 2019; **15**: e1007763.
- 8 Hawkey PM. The growing burden of antimicrobial resistance. *J Antimicrob Chemother* 2008; **62** Suppl 1: i1–i9.
- 9 Tacconelli E. Screening and isolation for infection control. *J Hosp Infect* 2009; **73**: 371–7.
- 10 Getoor L, Rhee JT, Koller D *et al*. Understanding tuberculosis epidemiology using structured statistical models. *Artif Intell Med* 2004; **30**: 233–56.
- 11 Koller D, Friedman N. *Probabilistic Graphical Models*. MIT Press, 2009.
- 12 Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol* 2008; **8**: 70.
- 13 Cherny SS, Nevo D, Baraz A *et al*. Revealing antibiotic cross-resistance patterns in hospitalized patients through Bayesian network modelling. *J Antimicrob Chemother* 2021; **76**: 239–48.
- 14 Love WJ, Zawack KA, Booth JG *et al*. Markov networks of collateral resistance: national antimicrobial resistance monitoring system surveillance results from *Escherichia coli* isolates, 2004–2012. *PLoS Comput Biol* 2016; **12**: e1005160.
- 15 Love WJ, Zawack KA, Booth JG *et al*. Phenotypical resistance correlation networks for 10 non-typhoidal *Salmonella* subpopulations in an active antimicrobial surveillance programme. *Epidemiol Infect* 2018; **146**: 991–100.
- 16 CDC Emerging Infections Program. Invasive *Staphylococcus aureus* Infection Tracking. <https://www.cdc.gov/hai/eip/saureus.html>.
- 17 Hernandez JM, Conforti P. Use of multivariate analysis to compare antimicrobial agents on the basis of *in vitro* activity data. *Antimicrob Agents Chemother* 1994; **38**: 184–8.

- 18 Tibshirani R. Regression shrinkage and selection via the lasso. *J R Statist Soc B* 1994; **58**: 267–88.
- 19 Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 2008; **9**: 432–41.
- 20 Lin J, Basu S, Banerjee M et al. Penalized maximum likelihood estimation of multi-layered Gaussian graphical models. *J Mach Learn Res* 2016; **17**: 1–51.
- 21 Javanmard A, Montanari A. Confidence intervals and hypothesis testing for high-dimensional regression. *J Mach Learn Res* 2014; **15**: 2869–909.
- 22 Liu H, Roeder K, Wasserman L. Stability approach to regularization selection (StARS) for high dimensional graphical models. *Adv Neural Inf Process Sys* 2010; **24**: 1432–40.
- 23 Core Team R. *R: A Language and Environment for Statistical Computing*. R Foundation For Statistical Computing, 2022.
- 24 Lin J. MultiLayerGGM. GitHub. <https://github.com/jhlinplus/MultiLayerGGM>.
- 25 Liu C, Bayer A, Cosgrove S et al. Clinical practice guidelines by the Infectious Diseases Society of America for the treatment of methicillin-resistant *Staphylococcus aureus* infections in adults and children. *Clin Infect Dis* 2011; **52**: 285–92.
- 26 Patel JB, Jevitt LA, Hageman J et al. An association between reduced susceptibility to daptomycin and reduced susceptibility to vancomycin in *Staphylococcus aureus*. *Clin Infect Dis* 2006; **42**: 1652–3.
- 27 Goldstein BP. Resistance to rifampicin: a review. *J Antibiot* 2014; **67**: 624–30.
- 28 Tran TT, Munita JM, Arias CA. Mechanisms of drug resistance: daptomycin resistance. *Ann NY Acad Sci* 2015; 1354: 32–53.
- 29 Seah C, Alexander DC, Louie L et al. MupB, a new high-level mupirocin resistance mechanism in *Staphylococcus aureus*. *Antimicrob Agents Chemother* 2016; **56**: 1916–20.
- 30 Patel JB, Gorwitz RJ, Jernigan JA. Mupirocin resistance. *Clin Infect Dis* 2009; **49**: 935–41.
- 31 Rahman M, Noble W, Cookson B. Transmissible mupirocin resistance in *Staphylococcus aureus*. *Epidemiol Infect* 1989; **102**: 261–70.
- 32 Morton TM, Johnston JL, Patterson J et al. Characterization of a conjugative staphylococcal mupirocin resistance plasmid. *Antimicrob Agents Chemother* 1995; **39**: 1272–80.
- 33 Kehrenberg C, Ojo K, Schwarz S. Nucleotide sequence and organization of the multiresistance plasmid pSCFS1 from *Staphylococcus sciuri*. *J Antimicrob Chemother* 2004; **54**: 936–9.
- 34 Knight GM, Budd EL, Whitney L et al. Shift in dominant hospital-associated methicillin-resistant *Staphylococcus aureus* (HA-MRSA) clones over time. *J Antimicrob Chemotherapy* 2012; **67**: 2514–22.
- 35 Blumberg HM, Rimland D, Carroll DJ et al. Rapid development of ciprofloxacin resistance in methicillin-susceptible and -resistant *Staphylococcus aureus*. *J Infect Dis* 1991; **163**: 1279–85.
- 36 Lakhundi S, Zhang K. Methicillin-resistant *Staphylococcus aureus*: molecular characterization, evolution, and epidemiology. *Clin Microbiol Rev* 2018; **31**: e00020-18.
- 37 See I, Mu Y, Albrecht V et al. Trends in incidence of methicillin-resistant *Staphylococcus aureus* bloodstream infections differ by strain type and healthcare exposure, United States, 2005–2013. *Clin Infect Dis* 2020; **71**: 19–25.
- 38 Andreatos N, Shehade F, Pliakos EE et al. The impact of antibiotic prescription rates on the incidence of MRSA bloodstream infections: a county-level, US-wide analysis. *Int J Antimicrob Agents* 2018; **52**: 195–200.
- 39 Jamrozy DM, Harris SR, Mohamed N et al. Pan-genomic perspective on the evolution of the *Staphylococcus aureus* USA300 epidemic. *Microb Genom* 2016; **2**: e000058.
- 40 Klevens RM, Morrison MA, Nadle J et al. Invasive methicillin-resistant *Staphylococcus aureus* infections in the United States. *JAMA* 2007; **298**: 1763–71.
- 41 Dantes R, Mu Y, Belflower R et al. National burden of invasive methicillin-resistant *Staphylococcus aureus* infections, United States, 2011. *JAMA Int Med* 2013; **173**: 1970–78.
- 42 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.
- 43 CLSI. *Performance Standards for Antimicrobial Disk Susceptibility Tests—Twenty-Ninth Edition: M100*. 2019.