Analysis of Multidrug Resistance in *Staphylococcus aureus* with a Machine Learning-Generated Antibigram

Casey L. Cazer,^a Lars F. Westblade,^{b,c} Matthew S. Simon,^c Reed Magleby,^c Mariana Castanheira,^d James G. Booth,^e Stephen G. Jenkins,^{b,c} Yrjö T. Gröhn^a

^aDepartment of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York, USA

^bDepartment of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, New York, USA

^cDivision of Infectious Diseases, Department of Medicine, Weill Cornell Medicine, New York, New York, USA

^dJMI Laboratories, North Liberty, Iowa, USA

^eDepartment of Statistics and Data Science, Cornell University, Ithaca, New York, USA

ABSTRACT Multidrug resistance (MDR) surveillance consists of reporting MDR prevalence and MDR phenotypes. Detailed knowledge of the specific associations underlying MDR patterns can allow antimicrobial stewardship programs to accurately identify clinically relevant resistance patterns. We applied machine learning and graphical networks to quantify and visualize associations between resistance traits in a set of 1,091 *Staphylococcus aureus* isolates collected from one New York hospital between 2008 and 2018. Antimicrobial susceptibility testing was performed using reference broth microdilution. The isolates were analyzed by year, methicillin susceptibility, and infection site. Association mining was used to identify resistance patterns that consisted of two or more individual antimicrobial resistance (AMR) traits and quantify the association among the individual resistance traits in each pattern. The resistance patterns captured the majority of the most common MDR phenotypes and reflected previously identified pairwise relationships between AMR traits in *S. aureus*. Associations between β -lactams and other antimicrobial classes (macrolides, lincosamides, and fluoroquinolones) were common, although the strength of the association among these antimicrobial classes varied by infection site and by methicillin susceptibility. Association mining identified associations between clinically important AMR traits, which could be further investigated for evidence of resistance coselection. For example, in skin and skin structure infections, clindamycin and tetracycline resistance occurred together 1.5 times more often than would be expected if they were independent from one another. Association mining efficiently discovered and quantified associations among resistance traits, allowing these associations to be compared between relevant subsets of isolates to identify and track clinically relevant MDR.

KEYWORDS antibiogram, association mining, machine learning, multidrug resistance, *Staphylococcus aureus*, antibiotic resistance

Multidrug resistance (MDR) conferred by genetic coresistance (i.e., linkage between resistance genes conferring phenotypic resistance to multiple agents) or cross-resistance (i.e., one gene conferring resistance to multiple agents) (1) limits or eliminates therapeutic antimicrobial choices. For example, hospital-associated *Staphylococcus aureus* can harbor *SCCmecII* (staphylococcal cassette chromosome [SCC] *mec* type II), which contains a macrolide and lincosamide resistance gene (cross-resistance from *ermA*) plus a methicillin resistance gene (coresistance from *mec* linked to *ermA*) (2). Consequently, associations between antimicrobial resistance (AMR) traits have been observed. Here, statistical associations between phenotypic AMR traits are indicated with asterisks. Statistical associations between phenotypic resistance traits may appear when there is genetic coresistance

Citation Cazer CL, Westblade LF, Simon MS, Magleby R, Castanheira M, Booth JG, Jenkins SG, Gröhn YT. 2021. Analysis of multidrug resistance in *Staphylococcus aureus* with a machine learning-generated antibiogram. *Antimicrob Agents Chemother* 65:e02132-20. <https://doi.org/10.1128/AAC.02132-20>.

Copyright © 2021 American Society for Microbiology. All Rights Reserved.

Address correspondence to Casey L. Cazer, clc248@cornell.edu.

Received 6 October 2020

Returned for modification 28 October 2020

Accepted 24 December 2020

Accepted manuscript posted online

11 January 2021

Published 18 March 2021

or cross-resistance. We use the expression “resistance trait” to refer to phenotypic resistance to an individual antimicrobial agent and “resistance pattern” to indicate a combination of resistance traits. Among *S. aureus* isolates, previously identified associations between phenotypic resistance traits include methicillin * clindamycin (CM), methicillin * erythromycin (ERY), fusidic acid (FUS) * CM, FUS * ERY, trimethoprim-sulfamethoxazole (SXT) * CM, and SXT * ERY (3).

Associations between two resistance traits can be quantified statistically with contingency tables (3) or correlation coefficients (1, 4, 5). However, genetic coresistance or cross-resistance frequently confers phenotypic resistance to more than two individual antimicrobial agents or classes (6). Therefore, new methods are required to analyze and track MDR phenotypes to account for higher-order (e.g., three-way, four-way, five-way, etc.) associations among resistance traits. Association mining is a machine learning method originally developed to identify relationships in consumer purchasing behaviors (7, 8). Association rules have been utilized in hospital nosocomial and AMR surveillance systems (9–14) and microbiology laboratory validation processes (15). Here, we used association mining to identify resistance patterns that manifest in bacterial isolates, measure their frequency, and quantify the statistical association among the resistance traits within the pattern to give insight into coresistance and cross-resistance. Other machine learning approaches to AMR analysis have focused on associations between resistance patterns and covariates, such as patient demographics or hospital departments (10, 13), rather than characterizing the relationship between the individual resistance traits. Mapping the associations between resistance traits can reveal the potential for AMR coselection during antimicrobial therapy and suggest antimicrobial use policies to slow the evolution of MDR (16).

Clinicians and antimicrobial stewardship programs can track resistance patterns to survey for changes in genetic linkages and coselection risks. Akin to using a standard antibiogram to monitor trends in the prevalence of AMR, resistance pattern analysis can be used to monitor associations among resistance traits and identify clinically significant resistance patterns. Importantly, unsupervised machine learning techniques, such as association mining, do not require an *a priori* specification of resistance patterns, because the algorithms can efficiently identify all resistance patterns present and select those that represent statistically significant associations. This ensures that novel or emerging MDR phenotypes can be detected. We use association mining to detect resistance patterns in 1,091 *S. aureus* isolates collected from one New York medical center between 2008 and 2018, quantify the prevalence of MDR (defined as resistance to three or more antimicrobial classes), and create graphical networks to highlight resistance pattern trends. We created a toolbox of adaptable techniques to estimate the statistical significance of the resistance patterns and address missing data that arises from changes in antimicrobial susceptibility testing (AST) panels. Missing or inconsistent AST data is a common issue in historical clinical data sets, and our methods maximize the utility of incomplete data for detecting changes in MDR over time. In this work, we demonstrate that association mining plus innovative graphical displays can be used to assess MDR in clinical or surveillance settings, and we identify resistance patterns in *S. aureus* that suggest that MDR coselection could occur during standard antimicrobial therapies.

RESULTS

The overall prevalence of resistance to 28 antimicrobials, comprising 13 classes, and their associated breakpoints and abbreviations, is given in Table 1. The prevalence of resistance in the studied subsets (e.g., year, methicillin susceptibility—i.e., methicillin-susceptible *S. aureus* [MSSA] and methicillin-resistance *S. aureus* [MRSA]—and infection site) is presented in Table S1 in the supplemental material. Overall, 39% of isolates were MRSA, 39.5% were MDR, and 13.3% were pansusceptible. Only susceptible isolates were observed for vancomycin (VAN) (1,091 susceptible/1,091 isolates tested), daptomycin (DAP) (1,091 susceptible/1,091 tested), tedizolid (TZD) (390 susceptible/

TABLE 1 Breakpoints and prevalence of resistance to the antimicrobials included in resistance pattern analysis of 1,091 *Staphylococcus aureus* isolates^a

Antimicrobial class	Antimicrobial agent (abbreviation)	Nonsusceptible breakpoint ($\mu\text{g/ml}$)	Breakpoint source	% nonsusceptible (no. tested)
β -Lactam	Ceftaroline (CPT)	>1 ^b	CLSI	2.5 (1,035)
	Ertapenem (ETP)	>2	FDA	17 (110)
	Oxacillin (OXA)	>2	CLSI	39 (1,091)
	Penicillin (PEN)	>0.12	CLSI	85 (943)
Fluoroquinolone	Ciprofloxacin (CIP)	>1	CLSI	39 (800)
	Levofloxacin (LEV)	>1	CLSI	38 (1,090)
	Moxifloxacin (MXF)	>0.5	CLSI	36 (1,028)
Macrolide	Azithromycin (AZI)	>2	CLSI	62 (144)
	Clarithromycin (CLM)	>2	CLSI	55 (56)
	Erythromycin (ERY)	>0.5	CLSI	58 (1,089)
	Telithromycin (TEL)	>1	CLSI	15 (537)
Lincosamide	Clindamycin (CM)	>0.5	CLSI	16 (1,085)
Lipoglycopeptide	Oritavancin (ORI)	>0.12	CLSI	0.19 (1,030)
	Telavancin (TLV)	>0.12	CLSI	9.7 (1,091)
Aminoglycoside	Gentamicin (GEN)	>4	CLSI	3.6 (1,091)
Tetracycline	Doxycycline (DOX)	>4	CLSI	0.87 (923)
	Minocycline (MIN)	>4	CLSI	0.34 (290)
	Omadacycline (OMC)	>0.5	FDA	2.2 (458)
	Tetracycline (TET)	>4	CLSI	6.5 (1,078)
	Tigecycline (TGC) ^c	>0.5	FDA	0 (1,088)
Folate pathway antagonists	Trimethoprim-Sulfamethoxazole (SXT)	>2/38	CLSI	2.5 (1,091)
Lipopeptide	Daptomycin (DAP)	>1	CLSI	0 (1,091)
Fusidane	Fusidic acid (FUS)	>1	EUCAST	0.7 (718)
Oxazolidinone	Linezolid (LZD)	>4	CLSI	0.092 (1,091)
	Tedizolid (TZD)	>0.5	CLSI	0 (390)
Streptogramin	Quinupristin-dalfopristin (QDA)	>1	CLSI	0.29 (345)
Glycopeptide	Teicoplanin (TEC)	>8	CLSI	0 (1,090)
	Vancomycin (VAN)	>2	CLSI	0 (1,091)

^aIsolates were classified as nonsusceptible if the MIC was greater than the nonsusceptible breakpoint (either intermediate or resistant). The proportion of nonsusceptible isolates is reported up to two significant digits. Breakpoints were preferentially sourced from the Clinical and Laboratory Standards Institute (CLSI) document M100, 28th edition (35), the U.S. Food and Drug Administration (FDA) (42, 43), and the European Committee on Antimicrobial Susceptibility Testing (EUCAST) (37).

^bThe 29th and 30th editions of CLSI document M100 (44, 45) revised the ceftaroline-intermediate category, which is included in the nonsusceptible range here, to a susceptible dose-dependent category.

^cTigecycline, a glycylcycline, was analyzed with other tetracycline-class agents.

390 tested), teicoplanin (TEC) (1,090 susceptible/1,090 tested), and tigecycline (TGC) (1,088 susceptible/1,088 tested) (see Table S1 in the supplemental material). There was resistance to the newer antimicrobials telavancin (9.7% nonsusceptible, TLV) and ceftaroline (2.5% nonsusceptible, CPT). The most common resistance phenotypes included resistance to β -lactams, macrolides, fluoroquinolones, and lincosamides (Table 2).

Statistical analysis of multidrug resistance. If resistance traits assorted independently, MDR would occur by chance when resistance traits occurred together in an isolate. In this situation, the prevalence of a specific MDR pattern would be the product of individual resistance trait prevalences (1). However, some resistance traits are likely not independent from each other due to cross-resistance and coresistance mechanisms, which will affect the prevalence of MDR. Overall, in this *S. aureus* isolate collection, MDR occurred less frequently than would be expected if all the individual resistance traits assorted independently (Fig. 1). Among all *S. aureus* isolates, pansusceptible and one-class-resistant isolates occurred more frequently than expected, two-class- and three-class-resistant isolates occurred less frequently than expected, and resistance to

TABLE 2 The 20 most common resistance phenotypes^a

Resistance phenotype	<i>n</i> ^b	Recovered resistance pattern ^c
PEN	184	—
Pansusceptible	145	—
ERY, PEN	91	Yes
CIP, ERY, LEV, MXF, OXA, PEN	85	Yes
ERY, OXA, PEN	31	Yes
CIP, CM, ERY, LEV, MXF, OXA, PEN, TEL	29	Yes
ERY	28	—
PEN, TLV	25	No
CIP, CM, ERY, LEV, MXF, OXA, PEN	18	Yes
OXA, PEN	16	Yes
ERY, LEV, MXF, OXA	15	Yes
AZI, ERY, PEN	13	Yes
CIP, LEV, MXF, OXA, PEN	13	Yes
CIP, ERY, LEV, MXF, PEN	13	Yes
CIP, ERY, LEV, MXF, OXA, PEN, TLV	13	Yes
AZI, ERY, LEV, OXA, PEN	11	Yes
AZI, ERY	9	Yes
AZI, ERY, LEV, MXF, OXA, PEN	9	Yes
CIP, LEV, MXF, PEN	8	Yes
ERY, PEN, TET	8	Yes

^aEach line represents one complete resistance phenotype, with isolates being intermediate or resistant to the listed antimicrobials (abbreviations are given in Table 1).

^bNumber of isolates with the indicated phenotype. The total number of isolates analyzed was 1,091.

^cRecovered resistance pattern indicates that an antimicrobial resistance pattern containing some or all of the resistance traits listed in the phenotype was recovered with association set mining. —, ineligible to be recovered with association set mining because at least two antimicrobials are required for a pattern to be identified with association set mining.

five or more classes occurred more frequently than expected (Fig. 1A). When the isolates are stratified by methicillin susceptibility, there are more pansusceptible MSSA isolates (Fig. 1C) and more non-MDR (i.e., resistant to ≤ 2 antimicrobial classes) MRSA isolates than expected if the individual resistance traits were independent (Fig. 1E). Correspondingly, there are fewer three-class-resistant MSSA isolates and fewer three-class- and four-class-resistant MRSA isolates than expected (Fig. 1C and E). There are more unique class resistance phenotypes than expected for six-class- and seven-class-resistant isolates (Fig. 1B). However, the unique phenotype distribution generally falls within the expected range for an assumption of independently assorting resistance traits when the isolates are stratified by methicillin resistance (Fig. 1D and F). Therefore, the shifted unique phenotype distribution observed over all isolates (Fig. 1B) is due to the proportion of MRSA isolates in the sample, which have a greater probability of diverse six-class and seven-class resistance phenotypes.

Association set mining. (i) Resistance pattern descriptive results. The impact of the expected cross-support ratio (eCSR) and conditional lift (cLift) filters on the number of resistance patterns, number of resistance traits per pattern, and number of classes per pattern is given in Table S2. The resistance patterns that passed both the eCSR and cLift filters are presented in Table S3. Across all analyses, the five most common antimicrobial resistance patterns, collapsed into antimicrobial classes, were β -lactam * fluoroquinolone * macrolide (134 patterns; average cLift = 5), fluoroquinolone * macrolide (112 patterns; average cLift = 3.5), β -lactam * fluoroquinolone (106 patterns; average cLift = 4), fluoroquinolone * fluoroquinolone (55 patterns; average cLift = 5.1), and β -lactam * macrolide (38 patterns; average cLift = 1.5). The average cLift indicates that β -lactam and fluoroquinolone resistance occur together, on average across all analyses, 4 times more frequently than would be expected if they were independent random variables. The strength and frequency of the patterns vary by data set; for example, β -lactam * fluoroquinolone

FIG 1 Distribution of MDR and unique AMR class phenotypes. (A, C, and E) Percentages of isolates with zero (pansusceptible) through seven antimicrobial class resistances. MDR is defined as resistance to 3 or more antimicrobial classes. (B, D, and F) Distribution of unique phenotypes, defined at the antimicrobial class level, across the class resistance categories (i.e., zero through seven class resistances) for all isolates (B), MSSA (D), and MRSA (F). Each gray line represents one of 100 simulated data sets with independent resistance traits. Red lines indicate the distribution in the *Staphylococcus aureus* data set for all isolates (A and B), MSSA (C and D), and MRSA (E and F).

patterns occur 21 times in pneumonia isolates with an average cLift of 4.8, and they occur 16 times in MRSA isolates with an average cLift of 1.3.

The top five resistance patterns with the greatest average cLift were gentamicin (GEN) * SXT * tetracycline (TET) (MRSA and PIHP [pneumonia in hospitalized patients]

FIG 2 Networks of resistance patterns for MSSA (A) and MRSA (B). Resistance patterns decomposed into nodes (antimicrobials) connected by edges (e.g., the pattern PEN * LEV * ERY decomposes into PEN * LEV, PEN * ERY, and LEV * ERY). Edge color represents the strength of the association (conditional lift) between the AMR traits, with darker edges representing a stronger association. Edge width represents the number of patterns containing the connected antimicrobials, with wider edges representing a more frequent association. Antimicrobial nodes are colored by antimicrobial class and labeled with antimicrobial names (see Table 1 for abbreviations). Node pie charts represent resistance prevalence: color, nonsusceptible; gray, susceptible; white, not tested.

analyses; average cLift=83), CM * telithromycin (TEL) * TET * TLV (SSSI [skin and skin structure infection] analysis; cLift=66), CPT * omadacycline (OMC) * SXT (MRSA; cLift=65), CM * TEL * TET (MSSA, SSSI, 2012 analyses; average cLift=39), ciprofloxacin (CIP) * CM * ERY * levofloxacin (LEV) * moxifloxacin (MXF) * oxacillin (OXA) (MRSA, SSSI, 2010 analyses; average cLift=25). Very high cLift values typically occur when the resistance traits in a given pattern have a low prevalence; therefore, these cLift values should be interpreted with caution and do not necessarily represent a higher level of significance than a smaller cLift (e.g., cLift=10). For example, the CPT * OMC * SXT pattern with a cLift of 65 occurred in only one MRSA isolate. Within MRSA isolates tested against all three antimicrobials, CPT resistance prevalence is 3.4%, OMC resistance prevalence is 4.5%, and SXT resistance prevalence is 5.7%. If these particular resistance traits were independent, it would be highly unlikely that an isolate would be resistant to all three agents (e.g., expected prevalence of 0.009%, or 1 of 11,358 isolates expected to be resistant to all three). Therefore, the high cLift value (i.e., high prevalence of resistance to all three traits compared to chance) supports a positive association between the three resistance traits. The CPT * OMC * SXT association should be confirmed in a larger isolate collection because of its low prevalence in this data set (1 isolate out of 1,091).

(ii) Resistance pattern networks. The resistance patterns are represented graphically as connections (i.e., edges) between antimicrobial agents in Fig. 2, 3, and 4. These graphical depictions reduce the multiway relationships in the antimicrobial resistance patterns to multiple 2-way connections between the resistance traits in the pattern.

FIG 3 Networks of resistance patterns for *Staphylococcus aureus* by infection site: bloodstream infections (BSI) (A), intra-abdominal infections (IAI) (B), pneumonia in hospitalized patients (PHIP) (C), and skin and skin structure infections (SSSI) (D). Resistance patterns decomposed into nodes (antimicrobials) connected by edges (e.g., the pattern PEN * LEV * ERY decomposes into PEN * LEV, PEN * ERY, and LEV * ERY). Edge color represents the strength of the association (conditional lift) between the AMR traits, with darker edges representing a stronger association. Edge width represents the number of patterns containing the connected antimicrobials, with wider edges representing a more frequent association. Antimicrobial nodes are colored by antimicrobial class and labeled with antimicrobial names (see Table 1 for abbreviations). Node pie charts represent resistance prevalence: color, nonsusceptible; gray, susceptible; white, not tested.

Subnetworks with 100% density, in which all possible edges are present among a group of antimicrobials (i.e., nodes), may represent a resistance pattern containing all of the antimicrobials plus patterns containing subsets of the antimicrobials. For example, in the MSSA graph (Fig. 2A), there is a 100% density 3-node network between CM * TEL * TET resistance, which reflects one resistance pattern containing all three resistance traits plus resistance patterns containing two of the three resistant traits (Table S3). However, it is possible to form a 100% density subnetwork from a combination of resistance patterns, each containing only a subset of the antimicrobials in the subnetwork. For example, in the MRSA antimicrobial class graph (Fig. 2B), there is a 100% density 4-node subnetwork between CPT, GEN, OMC, and SXT but there are no resistance patterns containing all four resistance traits (Table S3). This subnetwork is formed by a combination of six resistance patterns: (i) CPT * GEN, (ii) CPT * OMC, (iii) CPT * OMC * SXT, (iv) GEN * OMC, (v) GEN * SXT, and (vi) OMC * SXT.

Association mining recovered 120 antimicrobial resistance patterns, following eCSR and cLift filters, from MRSA isolates and only 10 patterns from MSSA isolates (Table S2). The MRSA resistance graph is denser than the MSSA resistance graph, and there is little edge overlap; only CM * TEL, CIP * LEV * MXF, azithromycin (AZI) * ERY, and clarithromycin (CLM) * ERY edges occur in both graphs (Fig. 2). The MRSA graph contains two subnetworks involving different β -lactam agents: CPT * GEN * OMC * SXT * TET and β -lactam (ertapenem [ETP], OXA, and penicillin [PEN]) * fluoroquinolone * lincosamide * macrolide. Conversely, the MSSA graph contains three distinct subnetworks: TLV * fluoroquinolones (CIP, LEV, and MXF), CM * TEL * TET, and a within-macrolide subnetwork (ERY * AZI * CLM).

Isolates from SSSI created a denser resistance network than isolates from other infection sites (Fig. 3). A subnetwork of β -lactam * fluoroquinolone * macrolide is found in SSSI, bloodstream infection (BSI), and PIHP isolates (Fig. 3A and D). This three-class subnetwork arises from two-agent and three-agent patterns, comprising 57 unique resistance patterns in SSSI isolates, 39 patterns in BSI isolates, and 28 patterns in PIHP isolates (Table S3), with the strongest association (average cLift) in SSSI isolates (Fig. 3D). There were sufficient SSSI isolates (≥ 30 isolates per subset) for a subset analysis of SSSI MRSA isolates and SSSI MSSA isolates, which revealed similar subnetworks of PEN * fluoroquinolones * macrolides, although the SSSI MRSA isolates created a denser network than the SSSI MSSA isolates because of resistance patterns involving the other β -lactams CPT, ETP, and OXA (Fig. S4). Resistance patterns involving clinically relevant antimicrobials include CM * TET (Fig. 3D), TEL * TET (Fig. 3D), SXT * TET (Fig. 3C), CPT * TLV (Fig. 3C), CPT * SXT (Fig. 3D), and CM * TLV (Fig. 3A and D).

The year analysis resistance networks (Fig. 4) were generally less dense than those from other analyses, likely because dividing data by 11 years results in smaller data sets. Three of the 11 years (2008, 2009, and 2010) contained a β -lactam * fluoroquinolone * macrolide * lipoglycopeptide (TLV) subnetwork (Fig. 4A to C), and four years (2011, 2014, 2015, and 2016 [Fig. 4D, E, G, H, and I]) contained the subnetwork without TLV, which is consistent with 100% susceptibility to TLV in all years except 2008 through 2010 (Table 1). The 2012 network also included TET and CM (Fig. 4E). Resistance to oritavancin (ORI), another lipoglycopeptide, was identified in 2016 (Table 1), although no resistance patterns containing ORI met the eCSR *P* value criterion.

The resistance pattern analysis utilized clinical breakpoints to determine susceptibility and nonsusceptibility. We performed a similar analysis with epidemiologic cutoff values (ECV) to assess associations between non-wild-type traits (Fig. S1 and S3). Antimicrobials without established ECV (ETP, PEN, TEL, TLV, and OMC) were excluded, and isolates that could not be classified as wild-type or non-wild-type because the AST dilution range did not span the ECV were treated as missing data. The ECV resistance pattern networks were generally less dense than the clinical breakpoint networks. The associations between non-wild-type traits were weaker (lighter line color) and less frequent (thinner line) than the same associations between resistance traits from the

FIG 4 Networks of resistance patterns for *Staphylococcus aureus* by year of collection for 2008 through 2018 (A through K). Resistance patterns decomposed into nodes (antimicrobials) connected by edges (e.g., the pattern PEN * LEV * ERY decomposes into PEN * LEV, PEN * ERY, and (Continued on next page)

clinical breakpoint analysis. The ECV analysis revealed an association between CM and CPT non-wild-type traits that was not found in the clinical breakpoint analysis (Fig. S2 and S3).

DISCUSSION

Staphylococcus aureus resistance dynamics are frequently driven by clonal shifts, which can be observed regionally and globally (2). For example, as the pulsed-field gel electrophoresis (PFGE) type USA300 replaced USA100, the prevalence of MDR MRSA decreased (17), because community-acquired MRSA isolates, including USA300, typically carry SCCs with fewer AMR genes than hospital-acquired MRSA isolates (2, 18, 19). Although data on SCCs or sequence types (ST) were not available for the 1,091 *S. aureus* isolates analyzed here, these clonal dynamics are likely responsible for the MDR trends observed. For example, the higher-than-expected prevalence of non-MDR MRSA (Fig. 1E) could be attributed to a predominance of a non-MDR clone in this medical center or the surrounding community. MSSA clone ST398 was circulating in Manhattan during this period (20) and was non-MDR (21). There were only five unique non-MDR MRSA phenotypes (at the class level), compared to 19 unique non-MDR MSSA phenotypes, which may reflect the number of unique circulating clones. For example, ST5 and ST8 clones accounted for 84% of MRSA BSIs in a New York medical center between 2007 and 2015 but only 30% of MSSA BSIs (22), suggesting greater diversity in the MSSA clone population. On the other hand, there is greater diversity in MDR MRSA isolates (29 unique class phenotypes and 64 three-class resistance patterns found by association mining) than MDR MSSA isolates (23 unique class phenotypes and 1 three-class resistance pattern found by association mining). There were six six-class (11 isolates) and two seven-class (4 isolates) unique phenotypes in MRSA isolates (Fig. 1F). This is more phenotype diversity than expected if all resistance traits assorted independently (Fig. 1F) and could also be attributed to clonal dynamics or the acquisition of resistance genes on top of the general background of β -lactam * fluoroquinolone * macrolide resistance.

Association mining can identify false-positive resistance patterns, which occur spuriously and do not represent true associations among resistance traits. True associations may arise from genetic mechanisms (i.e., coresistance and cross-resistance) or ecologic processes such as clonal dynamics and antimicrobial use trends. False-positive associations occur randomly and possibly in large numbers, depending on the size of data set analyzed. We aimed to reduce the false-positive rate by applying two separate tests of statistical association, the eCSR *P* value and the cLift bootstrap interval. These methods for testing the null hypothesis (namely, that there is no relationship between resistance traits) differentiate our association mining approach from other machine learning methods that have been used to study MDR, such as cluster analysis (23, 24). However, the processing pipeline, intended to reduce the risk of false positives, can create false negatives when a resistance pattern that represents a true association fails to pass the stringent eCSR *P* value and cLift bootstrap filters. For example, all isolates with ETP resistance are also resistant to PEN and OXA, but ETP is never directly connected to PEN and only sometimes connected to OXA (Fig. 3C). The resistance patterns that included PEN * OXA * ETP had small eCSR values because PEN resistance is very common compared to most other resistance traits (Table S1); hence, they were excluded from the final analysis because they did not pass the stringent eCSR *P* value filter. This exclusion is a false-negative result because there is a known association between these resistance traits: *mecA* confers resistance to all β -lactams except CPT unless specific point mutations are present (25).

The statistical power, sensitivity, and specificity of this analytic technique when

FIG 4 Legend (Continued)

LEV * ERY). Edge color represents the strength of the association (conditional lift) between the AMR traits, with darker edges representing a stronger association. Edge width represents the number of patterns containing the connected antimicrobials, with wider edges representing a more frequent association. Antimicrobial nodes are colored by antimicrobial class and labeled with antimicrobial names (see Table 1 for abbreviations). Node pie charts represent resistance prevalence: color, nonsusceptible; gray, susceptible; white, not tested.

applied to AST data are unknown and likely vary with each data set. Simulation studies or data sets with annotated genetic resistance mechanisms are required to estimate the sensitivity and specificity. However, a number of the resistance patterns identified by association mining were consistent with previously observed pairwise resistance associations and known genetic coresistance mechanisms in *S. aureus*, giving confidence that the resistance patterns capture true associations among resistance traits and not false-positive associations. To minimize the effect of false negatives, one can consider the relationships between antimicrobial classes rather than antimicrobial agents.

At the class level, we identified frequent associations between β -lactams, fluoroquinolones, macrolides, and, sometimes, lincosamides (Fig. 2 to 4). This is in agreement with previously identified pairwise associations between methicillin/OXA and ERY (3, 26), methicillin/OXA and CM (3, 26), and OXA and CIP (26). These specific agent associations, except for OXA * CM, were also frequently recovered by association mining (Fig. 2 to 4). β -Lactam, lincosamide, and macrolide resistance genes are found on *S. aureus* plasmids; macrolide resistance genes can also be found on some SCCmec types (19, 27). Fluoroquinolone resistance in *S. aureus* can result from chromosomal efflux pumps or point mutations in DNA gyrase or topoisomerase (28). An association between β -lactam resistance (primarily OXA), presumably mediated by mobile genetic elements, and fluoroquinolone resistance, presumably chromosomal, was driven by MRSA isolates. A frequent association between chromosomal resistance and mobile genetic element resistance could arise from one clone, or a few related clones, with a high prevalence in the sample population. Most likely, multiple β -lactam resistance genes circulate in this population of *S. aureus*, each with a different association with other resistance genes. This could explain the different associations observed for OXA and PEN resistance, which are associated with macrolide and fluoroquinolone resistance, and CPT resistance, which is associated with aminoglycoside, tetracycline, and sulfonamide resistance (Fig. 2 to 4). Genetic analysis of these isolates is required to confirm that β -lactam resistance genes (e.g., *mecA* and *mecA* mutants) are differentially linked with resistance genes of other antimicrobial classes through mobile genetic elements or clonal dynamics.

Association mining captures the cumulative effect of multiple similar phenotypes by identifying partial phenotypes (i.e., resistance patterns) that are consistent across several full phenotypes (e.g., the phenotypes ERY * CIP * GEN and ERY * CIP * TET share the resistance pattern ERY * CIP). This results in a more detailed understanding of MDR compared to other analysis methods, such as tabulating the most common resistance phenotypes. For example, ERY * PEN is found in 11/20 of the most common resistance phenotypes (Table 2) and is recovered in 9/17 of the resistance networks (Fig. 2 to 4). Association mining enables quantification of the relationship between ERY and PEN in each subset of isolates, accounting for all resistance patterns that contain the two agents. ERY and PEN resistance occur together 1.5 (average cLift) times more frequently than expected by chance in MRSA isolates, 1.7 times more frequently than expected in BSIs, and 1.5 times more frequently than expected in SSSI. Visualizations may be beneficial for antibiogram end users (e.g., clinicians, pharmacists, etc.) to gain a sense of trends over time. For example, the association between ERY and OXA weakens from 2008 (average cLift = 6) to 2015 (average cLift = 1.7), represented by a decrease in edge color intensity in Fig. 4H compared to Fig. 4A. This could indicate shifts in clonal prevalence or responses to antimicrobial stewardship.

Several resistance patterns involved the antimicrobials most commonly used to treat MRSA infections, suggesting a potential for coselection of AMR to multiple agents during routine clinical treatment. For example, uncomplicated SSSI with suspected or confirmed methicillin resistance are treated with CM, SXT, or a TET (29, 30). We uncovered an uncommon (4 isolates) but strong (cLift = 4.7) association between SXT and TET in MRSA isolates (Fig. 2B; also, see Table S3). Among all SSSI isolates, eight contained the CM * TET resistance pattern (cLift = 1.5); six of these isolates were also resistant to OXA. These associations suggest that treatment with TET could coselect for resistance to CM and/or SXT, or vice versa, resulting in a decline in the efficacy of two or

three clinically important antimicrobials concurrently. Clinical studies with repeated sampling and genetic sequencing would be required to demonstrate and quantify this coselection potential.

The association mining approach to an MDR-focused antibiogram goes beyond the information presented in a standard MDR analysis via tabulation (Table 2) to provide a quantitative analysis of all antimicrobial resistance patterns. The association sets identified resistance patterns from 16/20 of the most common resistance phenotypes (Table 2); three of the phenotypes contained ≤ 1 resistance trait, making them ineligible for association sets, which must contain at least two resistance traits. Therefore, this method captures the expected information from a standard MDR analysis but also can discover novel MDR patterns in a data set. The Apriori algorithm used by association mining is essential for analyzing a large number of AMR traits, because the number of potential resistance patterns grows exponentially with the number of antimicrobials included in the AST panel. This data set with 28 antimicrobials could yield up to 268,435,455 resistance patterns ($2^k - 1$, where k is the number of antimicrobials).

Resistance pattern quantification and visualization methods bring additional value beyond simple MDR tabulation methods and permit comparisons across groups or time periods. We find that compared to previous applications of association mining (31), association sets, rather than association rules, are easier to interpret as relationships among resistance traits. Current antibiogram approaches (17, 32, 33) ignore antimicrobial agents with substantial amounts of missing AST results due to AST panel variation, impeding detection of novel or emerging MDR patterns. We created quality metrics (e.g., eCSR and cLift) to accommodate missing AST results, which occurs when clinical laboratories change AST panels to address shifts in clinical practice and antimicrobial formularies. The association mining antibiogram describes the strength and frequency of a resistance pattern with lift, CSR, support, and their derivatives, offering new metrics for tracking resistance patterns, which were previously described by prevalence only (Table 2). The statistical significance of each resistance pattern can be assessed with statistical simulations, as demonstrated here with bootstrapped confidence intervals and P value calculations. Our metrics to address bias from missing data and our tools for statistical inference (see the supplemental methods) advance MDR epidemiology by creating a toolbox for analyzing clinical AST data sets with association mining. Additional theoretical research on applying hypothesis tests, controlling the false discovery rate, and accounting for missing data in association sets will improve MDR analysis and address some of the limitations of this method. Quantitative association mining, in which nominal or ordinal levels of each variable are considered (e.g., MICs) rather than a binary response (e.g., resistant or susceptible), could increase our understanding of resistance trait associations, but applying this method to AMR data requires further investigation.

Association mining is a flexible tool and could incorporate STs, clonal complexes, and other variables when data are available. For example, ST can be included in association sets or used to develop association rules of the form $\{ST\} \rightarrow \{AMR\}$, in which $\{ \}$ indicates a set of variables and $\rightarrow$ indicates a relationship between the sets. An expanded set of quality metrics can be used to quantify the relationship between ST and resistance patterns. If AST data can be combined with individual-level antimicrobial use (AMU) data, association rules of the form $\{AMU\} \rightarrow \{AMR\}$ can be generated to investigate the impact of prescribing on MDR patterns. Aggregate AMU data or population-level antimicrobial stewardship information can be used to guide comparisons of resistance patterns across time or hospital units to understand the emergence of clinically relevant resistance patterns. Similarly, genomic data could also be used to build association rules of the form $\{\text{genes}\} \rightarrow \{\text{AMR phenotype}\}$, which could be used to uncover novel AMR genes or relationships between reduced susceptibility (i.e., specific MIC values or ranges) and specific genes or gene expression. Phenotypic resistance patterns could be considered an aggregate of the underlying genotypic resistance patterns, which can be identified by mining association sets from a data set of AMR genes or

whole genomic sequences. Therefore, historic phenotypic AMR data and the increasing amounts of genomic AMR data can be used concurrently to track MDR patterns across time.

In conclusion, we propose association mining as a rigorous and effective method to assess and monitor MDR in clinical and surveillance settings. We have demonstrated that association mining can identify resistance patterns consistent with common phenotypes but provides substantial additional information on the strength and frequency of these associations. Resistance patterns are also easily visualized by decomposing the patterns into networks of nodes (antimicrobials) and edges (associations between antimicrobials). Association mining identified resistance patterns in *S. aureus* from one New York hospital that were congruent with previously identified pairwise resistance associations and genetic coresistance mechanisms. Furthermore, problematic associations between antimicrobials used to treat *S. aureus* can be measured and tracked with association mining and network visualization. Therefore, adding this method to a traditional antibiogram analysis substantially improves our understanding of resistance trait associations and MDR. Future work could include exploring the impact of these associations on coselection and clinical antimicrobial use with clinical studies, further refining association set mining with hypothesis testing and power analysis methods, expanding the method to Gram-negative organisms, and developing interactive applications for clinicians to query and visualize a machine learning antibiogram.

MATERIALS AND METHODS

Isolate collection and antimicrobial susceptibility testing. Since 1997, hospitals in North America, Europe, Latin America, and the Asia-Pacific region have submitted bacterial isolates to the SENTRY Antimicrobial Surveillance Program. Each year, participating institutions consecutively collect isolates by infection site, and the antimicrobial susceptibility profile of these isolates is determined at JMI Laboratories (North Liberty, IA). All isolates included in this study were collected at one hospital in New York between 2008 and 2018. Isolates from bloodstream infections (BSIs), pneumonia in hospitalized patients (PIHP), skin and skin structure infections (SSSI), and intra-abdominal infections (IAIs) were assayed against a panel of antimicrobial agents (Table 1) using reference broth microdilution (34). A core panel of antimicrobial agents were used for almost every isolate: CM, DAP, ERY, GEN, LEV, linezolid (LZD), OXA (to test for resistance to methicillin), SXT, TEC, TET, TGC, TLV, and VAN. Other antimicrobial agents were tested for some isolates based on antimicrobial availability, clinical use, and stakeholder interest. Quality control (QC) was performed using recommended QC isolates and procedures. Species identification was confirmed, when necessary, using biochemical assays or by matrix-assisted laser desorption/ionization–time of flight mass spectrometry (Bruker Daltonics, Inc., Billerica, MA), following manufacturer instructions.

Statistical analysis of multidrug resistance. MICs were interpreted as susceptible or nonsusceptible (intermediate or resistant) based on Clinical and Laboratory Standards Institute (CLSI) breakpoints (35) (Table 1). U.S. Food and Drug Administration (36) and European Committee on Antimicrobial Susceptibility Testing (37) breakpoints were also considered for antimicrobials not reported by CLSI (Table 1). Isolates were categorized for three separate analyses: (i) by year, (ii) by methicillin susceptibility (i.e., MSSA versus MRSA), and (iii) by infection site. Only infection site subsets with ≥ 30 isolates were included in the analysis; unspecified infection sites (e.g., “other”) were excluded.

The significance of MDR prevalence and MDR phenotypes was assessed with statistical simulation using a null hypothesis of no associations among resistance traits. We generated 100 “no-association” data sets with the same dimensions, resistance prevalence, and AST panels as the data set but with all resistance traits independent from each other (i.e., independent binomial random variables). We determined the null distribution of MDR prevalence and unique MDR phenotypes, at the antimicrobial class level, by calculating the number of isolates resistant to multiple antimicrobial classes and the number of unique resistance phenotypes in the “no-association” data sets. The multiclass resistance prevalence and multiclass resistance phenotypes in the observed data were compared to the null distribution to determine the likelihood that the observed data would occur under the null hypothesis of independent resistance traits. Additional discussion of the null hypothesis is presented in the supplemental methods.

Association set mining. The application of association mining to antimicrobial susceptibility data was described previously (31), and substantial details, including an example analysis, are provided in the supplemental methods. The analysis presented here utilizes association sets, also called frequent sets, rather than association rules (31), in which the associations between antimicrobials may be asymmetric. Using symmetric association sets improves the interpretability and visualization of resistance patterns. Here, we also used novel methods to address missing antimicrobial susceptibility data and perform statistical inference on the AMR trait associations. In brief, association set mining employs the Apriori algorithm (8), an unsupervised machine learning algorithm, to identify patterns in data sets containing binary variables. In the case of antimicrobial susceptibility data, the algorithm identifies combinations of resistance traits (referred to as patterns) that exist within isolates. We refer to the Centers for Disease Control

and Prevention (CDC) definition of resistance pattern, “a description of the antibiotic resistance testing results for an isolate” (38), and consider a resistance pattern to be any combination of resistance testing results or traits that occur within an isolate. The resistance pattern may include some or all of the antimicrobials in a tested panel (see the supplemental methods for an example). Resistance patterns are described by listing the antimicrobial agents or classes in the pattern connected by asterisks. The Apriori algorithm identified all resistance patterns involving two or more antimicrobials that occurred in at least one isolate.

For each analysis, data were divided by the relevant category (year, methicillin susceptibility [as determined by oxacillin susceptibility], and infection site), referred to as data sets, and resistance patterns were mined from each data set (e.g., each year, MSSA isolates, MRSA isolates, and each infection site). The frequency and association strength of each resistance pattern were assessed with three quality metrics: support (equation 1), cross-support ratio (equation 2), and lift (equation 3).

$$\text{Support}(X) = \frac{\text{count}(X)}{N} = P(X) \quad (1)$$

$$\text{CSR}(X) = \frac{\min_{x \in X}(\text{Support}(x))}{\max_{x \in X}(\text{Support}(x))} \quad (2)$$

$$\text{Lift}(X) = \frac{\text{Support}(X)}{\prod_{x \in X} \text{Support}(x)} \quad (3)$$

Support is the prevalence of the resistance pattern in the data set. The cross-support ratio compares the prevalence of individual resistance traits within the resistance pattern and ranges from 0 to 1. When all the resistance traits within a resistance pattern occur at approximately the same prevalence, the cross-support ratio is close to 1 and indicates a higher likelihood of a positive association among the resistance traits. If the prevalences of individual resistance traits are substantially different, the resistance pattern is more likely to be spurious, and thus, the cross-support ratio is close to 0 (39, 40). Lift, the ratio of observed to expected prevalence, ranges from 0 to ∞ and compares the prevalence of the resistance pattern in the data set to the expected prevalence of the resistance pattern under a null hypothesis of no association among resistance traits. A lift greater than 1 indicates that the resistance pattern occurs more frequently than expected under the null hypothesis, suggesting a positive association among the resistance traits within the pattern. A lift of 1 suggests no association among the resistance traits, and a lift of less than 1 indicates a negative association among the resistance traits (e.g., resistance to one antimicrobial is associated with susceptibility to another antimicrobial). Missing data can bias these quality measures. Therefore, we used two derivatives to account for changes in the composition of antimicrobial susceptibility panels used for AST: expected cross-support ratio (eCSR) and conditional lift (cLift). Details of these quality measure derivatives, including example calculations and interpretations, are given in the supplemental methods; eCSR and cLift are interpreted in the same way as cross-support ratio and lift, respectively.

where N is the number of isolates, X is the resistance pattern, and x is the resistance trait in the pattern.

Depending on data dimensions and resistance prevalence, association mining can identify thousands or tens of thousands of resistance patterns. Statistical significance of the patterns is challenging to establish, because the number of tested hypotheses exceeds the number of patterns identified (41). Therefore, we used two complementary methods, simulation and bootstrapping, to establish the significance of each pattern, using a null hypothesis of independent resistances (see the supplemental methods for details). First, we used a simulation, performed independently for each tested data set (e.g., each year, MSSA versus MRSA, or each infection site), to estimate the expected distribution of eCSR among spurious resistance patterns with no true association among resistance traits (referred to as eCSR null distribution). The one-sided P value for the eCSR of each resistance pattern in the tested data set was calculated by comparing the eCSR of the pattern to the eCSR null distribution. This one-sided test selects for a large eCSR, which represents a positive association between resistance traits. Resistance patterns with an eCSR P value of ≤ 0.05 were further filtered using cLift. A 95% confidence interval on the cLift for each resistance pattern was determined with 1,000 bootstrap resamples with a size of $r - 1$ (r = size of the data set analyzed) and the 2.5% and 97.5% bootstrap percentiles. Resistance patterns with an eCSR P value of ≤ 0.05 and a cLift bootstrap interval that excluded 1 were selected for further analysis.

The remaining resistance patterns (following eCSR and cLift filters) were visualized as a network by decomposing each pattern into all possible pairwise combinations of antimicrobials (nodes) within the pattern, and each pair was connected with an edge (line). For example, the resistance pattern with ampicillin (AMP), CIP, and AZI resistance traits decomposes into three pairwise combinations: (i) AMP * CIP, (ii) AMP * AZI, and (iii) CIP * AZI. Edges were summed across all resistance patterns such that edge width was proportional to the number of resistance patterns that contained the edge (i.e., the number of patterns that contained the two antimicrobials connected by the edge), and the edge color corresponded to the average cLift of the represented patterns (see the supplemental methods for details and an example). Resistance pattern networks were created for each analysis (by year, by methicillin susceptibility [MSSA and MRSA], and by infection site).

Data availability. All analyses were implemented in R (version 3.6.1); all R code necessary to replicate the results of this study is publicly available (<https://doi.org/10.5281/zenodo.3887072>), and data are available upon request (<https://doi.org/10.5281/zenodo.3887031>).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 1.3 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.05 MB.

SUPPLEMENTAL FILE 3, XLSX file, 0.03 MB.

ACKNOWLEDGMENTS

C.L.C. was supported by the Office of the Director, National Institutes of Health, under award number T32OD011000.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

C.L.C.: Conceptualization, Formal analysis, Methodology, Software, Interpretation of results, Validation, Visualization, Writing – original draft, review & editing. L.F.W.: Conceptualization, Investigation, Resources, Interpretation of results, Writing – review & editing. M.S.S.: Conceptualization, Interpretation of results, Writing – review & editing. R.M.: Conceptualization, Interpretation of results, Writing – review & editing. M.C.: Data curation, Investigation, Methodology, Interpretation of results, Project administration (SENTRY), Resources, Writing – review & editing. J.G.B.: Methodology, Supervision, Writing – review & editing. S.G.J.: Data curation, Investigation, Resources, Writing – review & editing. Y.T.G.: Conceptualization, Resources, Supervision, Writing – review & editing.

REFERENCES

- Chang HH, Cohen T, Grad YH, Hanage WP, O'Brien TF, Lipsitch M. 2015. Origin and proliferation of multiple-drug resistance in bacterial pathogens. *Microbiol Mol Biol Rev* 79:101–116. <https://doi.org/10.1128/MMBR.00039-14>.
- Chambers HF, Deleo FR. 2009. Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat Rev Microbiol* 7:629–641. <https://doi.org/10.1038/nrmicro2200>.
- Wimmerstedt A, Kahlmeter G. 2008. Associated antimicrobial resistance in *Escherichia coli*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Streptococcus pneumoniae* and *Streptococcus pyogenes*. *Clin Microbiol Infect* 14:315–321. <https://doi.org/10.1111/j.1469-0691.2007.01946.x>.
- Love WJ, Zawack KA, Booth JG, Grohn YT, Lanzas C. 2018. Phenotypic resistance correlation networks for 10 non-typhoidal *Salmonella* subpopulations in an active antimicrobial surveillance programme. *Epidemiol Infect* 146:1–12. <https://doi.org/10.1017/S0950268818000833>.
- Love WJ, Zawack KA, Booth JG, Grün YT, Lanzas C. 2016. Markov networks of collateral resistance: national antimicrobial resistance monitoring system surveillance results from *Escherichia coli* isolates, 2004–2012. *PLoS Comput Biol* 12:e1005160. <https://doi.org/10.1371/journal.pcbi.1005160>.
- Costa SS, Viveiros M, Amaral L, Couto I. 2013. Multidrug efflux pumps in *Staphylococcus aureus*: an update. *Open Microbiol J* 7:59–71. <https://doi.org/10.2174/1874285801307010059>.
- Tan P-N, Steinbach M, Karpatsch A, Kumar V. 2018. Association analysis: basic concepts and algorithms, introduction to data mining, 2nd ed. Pearson Education, New York, NY.
- Agrawal R, Imieliński T, Swami A. 1993. Mining association rules between sets of items in large databases, ACM SIGMOD record. ACM, Washington, DC. <https://doi.org/10.1145/170035.170072>.
- Giannopoulou EG, Kemerlis VP, Polemis M, Papaparaskevas J, Vatsopoulos AC, Vazirgiannis M. 2007. A large scale data mining approach to antibiotic resistance surveillance, p 439–444. In *Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07)*, Maribor, Slovenia, June 20–22 2007. IEEE, New York, NY. <https://doi.org/10.1109/cbms.2007.8>.
- Gerontini M, Vazirgiannis M, Vatsopoulos AC, Polemis M. 2011. Predictions in antibiotics resistance and nosocomial infections monitoring, p 1–6. In *24th International Symposium on Computer-Based Medical Systems (CBMS)*, Bristol, UK, June 27–30 2011. IEEE, New York, NY. <https://doi.org/10.1109/cbms.2011.5999112>.
- Ma L, Tsui F-C, Hogan WR, Wagner MM, Ma H. 2003. A framework for infection control surveillance using association rules, p 410–414. In *AMIA Annual Symposium Proceedings*. AMIA, Bethesda, MD.
- Brossette SE, Sprague AP, Hardin JM, Waites KB, Jones WT, Moser SA. 1998. Association rules and data mining in hospital infection control and public health surveillance. *J Am Med Inform Assoc* 5:373–381. <https://doi.org/10.1136/jamia.1998.0050373>.
- Jones WT, Sprague AP, Brossette SE, Moser SA. 2000. A data mining system for infection control surveillance. *Methods Infect Med* 39:303–310. <https://doi.org/10.1055/s-0038-1634449>.
- Moser SA, Jones WT, Brossette SE. 1999. Application of data mining to intensive care unit microbiologic data. *Emerg Infect Dis* 5:454–457. <https://doi.org/10.3201/eid0503.990320>.
- Lamma E, Manservigi M, Mello P, Nanetti A, Riguzzi F, Storari S. 2001. The automatic discovery of alarm rules for the validation of microbiological data. *Proceedings of Workshop on Intelligent Data Analysis in Medicine and Pharmacology*, London, UK (IDAMAP).
- Pal C, Papp B, Lazar V. 2015. Collateral sensitivity of antibiotic-resistant microbes. *Trends Microbiol* 23:401–407. <https://doi.org/10.1016/j.tim.2015.02.009>.
- Diekema DJ, Pfaller MA, Shortridge D, Zervos M, Jones RN. 2019. Twenty-year trends in antimicrobial susceptibilities among *Staphylococcus aureus* from the SENTRY Antimicrobial Surveillance Program. *Open Forum Infect Dis* 6:S47–S53. <https://doi.org/10.1093/ofid/ofy270>.
- Diekema DJ, Richter SS, Heilmann KP, Dohm CL, Riahi F, Tendolkar S, McDanel JS, Doern GV. 2014. Continued emergence of USA300 methicillin-resistant *Staphylococcus aureus* in the United States: results from a nationwide surveillance study. *Infect Control Hosp Epidemiol* 35:285–292. <https://doi.org/10.1086/675283>.
- Malachowa N, DeLeo FR. 2010. Mobile genetic elements of *Staphylococcus aureus*. *Cell Mol Life Sci* 67:3057–3071. <https://doi.org/10.1007/s00018-010-0389-4>.
- Uhlemann AC, Hafer C, Miko BA, Sowash MG, Sullivan SB, Shu Q, Lowy FD. 2013. Emergence of sequence type 398 as a community- and healthcare-associated methicillin-susceptible staphylococcus aureus in northern Manhattan. *Clin Infect Dis* 57:700–703. <https://doi.org/10.1093/cid/cit375>.
- Mediavilla JR, Chen L, Uhlemann A-C, Hanson BM, Rosenthal M, Stanak K, Koll B, Fries BC, Armellino D, Schilling ME, Weiss D, Smith TC, Lowy FD, Kreiswirth BN. 2012. Methicillin-susceptible *Staphylococcus aureus* ST398, New York and New Jersey, USA. *Emerg Infect Dis* 18:700–702. <https://doi.org/10.3201/eid1804.111419>.
- Austin ED, Sullivan SS, Maccesic N, Mehta M, Miko BA, Nematollahi S, Shi Q, Lowy FD, Uhlemann AC. 2020. Reduced mortality of *Staphylococcus aureus* bacteremia in a retrospective cohort study of 2139 patients: 2007–2015. *Clin Infect Dis* 70:1666–1674. <https://doi.org/10.1093/cid/ciz498>.

23. Jahne MA, Rogers SW, Ramler IP, Holder E, Hayes G. 2015. Hierarchical clustering yields insight into multidrug-resistant bacteria isolated from a cattle feedlot wastewater treatment system. *Environ Monit Assess* 187:4168. <https://doi.org/10.1007/s10661-014-4168-9>.
24. Coelho JR, Carrico JA, Knight D, Martinez JL, Morrissey I, Oggioni MR, Freitas AT. 2013. The use of machine learning methodologies to analyse antibiotic and biocide susceptibility in *Staphylococcus aureus*. *PLoS One* 8:e55582. <https://doi.org/10.1371/journal.pone.0055582>.
25. Chan LC, Basuino L, Diep B, Hamilton S, Chatterjee SS, Chambers HF. 2015. Ceftobiprole- and ceftaroline-resistant methicillin-resistant *Staphylococcus aureus*. *Antimicrob Agents Chemother* 59:2960–2963. <https://doi.org/10.1128/AAC.05004-14>.
26. Fluit AC, Schmitz FJ, Verhoef J, European SENTRY Participants. 2001. Multi-resistance to antimicrobial agents for the ten most frequently isolated bacterial pathogens. *Int J Antimicrob Agents* 18:147–160. [https://doi.org/10.1016/s0924-8579\(01\)00357-0](https://doi.org/10.1016/s0924-8579(01)00357-0).
27. Liu J, Chen D, Peters BM, Li L, Li B, Xu Z, Shirliff ME. 2016. Staphylococcal chromosomal cassettes *mec* (SCCmec): a mobile genetic element in methicillin-resistant *Staphylococcus aureus*. *Microb Pathog* 101:56–67. <https://doi.org/10.1016/j.micpath.2016.10.028>.
28. Kaatz GW, Seo SM. 1997. Mechanisms of fluoroquinolone resistance in genetically related strains of *Staphylococcus aureus*. *Antimicrob Agents Chemother* 41:2733–2737. <https://doi.org/10.1128/AAC.41.12.2733>.
29. David MZ, Daum RS. 2017. Treatment of *Staphylococcus aureus* infections, p 325–383. In Bagnoli F, Rappuoli R, Grandi G (ed), *Staphylococcus aureus: microbiology, pathology, immunology, therapy and prophylaxis* Springer International Publishing, Berlin, Germany.
30. Tong SY, Davis JS, Eichenberger E, Holland TL, Fowler VG, Jr. 2015. *Staphylococcus aureus* infections: epidemiology, pathophysiology, clinical manifestations, and management. *Clin Microbiol Rev* 28:603–661. <https://doi.org/10.1128/CMR.00134-14>.
31. Cazer CL, Al-Mamun MA, Kaniyamattam K, Love WJ, Booth JG, Lanzas C, Grohn YT. 2019. Shared multidrug resistance patterns in chicken-associated *Escherichia coli* identified by association rule mining. *Front Microbiol* 10:687. <https://doi.org/10.3389/fmicb.2019.00687>.
32. Cummings KJ, Aprea VA, Altier C. 2015. Antimicrobial resistance trends among canine *Escherichia coli* isolates obtained from clinical samples in the northeastern USA, 2004–2011. *Can Vet J* 56:393–398.
33. Hindler JF, Stelling J. 2007. Analysis and presentation of cumulative anti-diagrams: a new consensus guideline from the Clinical and Laboratory Standards Institute. *Clin Infect Dis* 44:867–873. <https://doi.org/10.1086/511864>.
34. Clinical and Laboratory Standards Institute. 2018. Methods for dilution antimicrobial susceptibility tests for bacteria that grow aerobically, 11th ed. Clinical and Laboratory Standards Institute, Wayne, PA.
35. Clinical and Laboratory Standards Institute. 2018. Performance standards for antimicrobial susceptibility testing, 28th ed. Clinical and Laboratory Standards Institute, Wayne, PA.
36. U.S. Food and Drug Administration. 17 June 2019. Antibacterial susceptibility test interpretive criteria. <https://www.fda.gov/drugs/development-resources/antibacterial-susceptibility-test-interpretive-criteria>. Accessed 25 July 2019.
37. European Committee on Antimicrobial Susceptibility Testing. 15 May 2018. Breakpoint tables for interpretation of MICs and zone diameters. <http://www.eucast.org>. Accessed 17 December 2018.
38. Centers for Disease Control and Prevention. 15 March 2019. Glossary of terms related to antibiotic resistance. <https://www.cdc.gov/narms/resources/glossary.html>. Accessed 28 January 2020.
39. Xiong H, Tan PN, Vipin K. 2003. Mining strong affinity association patterns in data sets with skewed support distribution. Proceedings of the Third IEEE International Conference on Data Mining, Melbourne, FL. IEEE, New York, NY.
40. Hahsler M. A probabilistic comparison of commonly used interest measures for association rules. http://michael.hahsler.net/research/association_rules/measures.html. Accessed 7 May 2018.
41. Megiddo N, Srikant R. 1998. Discovering predictive association rules, Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining, New York, NY. AAAI Press, Palo Alto, CA.
42. FDA. 17 June 2019. Tigecycline–injection products. <https://www.fda.gov/drugs/development-resources/tigecycline-injection-products>. Accessed 27 January 2020.
43. FDA. 13 June 2018. Ertapenem–injection products. <https://www.fda.gov/drugs/development-resources/ertapenem-injection-products>. Accessed 14 March 2019.
44. Clinical and Laboratory Standards Institute. 2020. Performance standards for antimicrobial susceptibility testing, 30th ed. Clinical and Laboratory Standards Institute, Wayne, PA.
45. Clinical and Laboratory Standards Institute. 2019. Performance standards for antimicrobial susceptibility testing, 29th ed. Clinical and Laboratory Standards Institute, Wayne, PA.